

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO,

THIS PAGE BLANK (USPTO)

to proteins encoded by genes of the human genome, useful in an assay for screening and identifying of one or more novel peptides which are drug candidates or pro-drugs -

```

Query Match      91.7%; Score 22; DB 22; Length 10;
Best Local Similarity 66.7%; Pred. NO. 47;
Matches 4; Conservative 0; Mismatches 2; Indels 0;
Gaps 0;
1 gtxxps 6
5 atssos 10

```

NSULT 2
G97971 Human complementary peptide, SEQ ID NO: 4166.
AAG97971 standard; Peptide; 10 AA.
AAG97971; Homo sapiens.
18-SEP-2001 (first entry)
Human complementary peptide, SEQ ID NO: 4166.
Human; complementary peptide; ligand; drug discovery; drug design.
Homo sapiens.
18-SEP-2001 (first entry)
Human complementary peptide, SEQ ID NO: 4166.
Human; complementary peptide; ligand; drug discovery; drug design.
Homo sapiens.

14 - JUN - 2001 .
13 - DEC - 2000 ; 2000WO-GB04776 .
13 - DEC -1999 ; 99GB-0029464 .

Roberts GW, Heal JR;
WPI: 2001-408419/43.
A set of peptide ligands consisting of specific complementary peptides to proteins encoded by genes of the human genome, useful in an assay

Example 6: Page 630; 646pp; English.
The invention relates to a set of complementary peptide ligands generated from the human genome. The complementary peptides interact with their relevant target proteins encoded in the human genome. They can be used as reagents in drug discovery and as lead drug candidates or prodrugs.

Sequence is a complementary peptide provided in the specification.

QY 1 gtxxps 6
Db 5 qssps 10

The invention relates to a set of complementary peptide ligands generated from the human genome. The complementary peptides interact with their relevant target proteins encoded in the human genome. They can be used as reagents in drug discovery and as lead ligands to facilitate drug design and development. The present sequence is a complementary peptide provided in the specification.

Sequence	10 AA;	Sequence	10 AA;
Query Match	91.7%	Score 22;	DB 22;
Best Local Similarity	66.7%;	Pred. No. 47;	Length 10;
Matches 4;	Conservative	0;	Mismatches
		2;	Indels
		0;	
1	atcccccc	6	

5 gtssps 10
RESULTS 2
G97971 2
AAG97971 standard; Peptide; 10 AA.

18-SEP-2001 (first entry)
Human complementary peptide, SEQ ID NO: 4166.
Human; complementary peptide; ligand; drug discovery; drug design.
Homo sapiens.

WO200422774A2.
14 - JUN - 2001.
13 - DEC - 2000; 2000WO - GB04776.
13 - DEC - 1999; 99GB - 0029464.

Roberts GW, Heal JR;
WPI; 2001-408419/43.

A set of peptide ligands consisting of specific complementary peptides to proteins encoded by genes of the human genome, useful in an assay for screening and indentifying one or more novel peptides which are

Example 6: Page 630: 646pp; English.
The invention relates to a set of complementary peptide ligands generated from the human genome. The complementary peptides interact with their relevant target proteins encoded in the human genome. They can be used as reagents in drug discovery and as lead compounds in drug development.

Sequence	10 AA:	Query Match	91.7%	Score 22;	DB 22;	Length 10;
Best Local Similarity	66.7%	Pred. No.	47;			

QY 1 gtxxps 6
Db 5 qssps 10

RESULT 3
 AAG97972
 ID AAG97972 standard; Peptide; 10 AA.
 XX
 AC
 AAG97972;
 XX

XX DE Human complementary peptide, SEQ ID NO: 4167.
XX KW Human; complementary peptide; ligand; drug discovery; drug design.
XX OS Homo sapiens.
XX

PD 14 - JUN - 2001.
XX
PF 13 - DEC - 2000; 2000WO - GB04776.
XX
PR 13 - DEC - 1999; 99GB - 0029464.
XX
PA (PROT-) PROTEOM LTD.

PS Example 6; page 630; 646pp; English.
XX CC the invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
CC

SQ	Sequence	10 AA:
Query	Match	91.7%
Best Local	Similarity	66.7%
Matches	4	
Conervative	0	
Matches	4	
Mismatches	0	
Indels	2	
Gaps	0	
Length	10	

RESULT 4
AAG97973
ID AAG97973 standard; Peptide; 10 AA.
xx

18-SEP-2001 (first entry)
DT Human complementary peptide, SEQ ID NO: 4168.
XX Human; complementary peptide; ligand; drug discovery; drug design.
XX KW Homo sapiens.
XX OS

XX
 PN Example 6; Page 630; 646pp; English.
 XX
 XX
 CC the invention relates to a set of complementary peptide ligands
 CC generated from the human genome. The complementary peptides
 CC interact with their relevant target proteins encoded in the human
 CC genome. They can be used as reagents in drug discovery and as lead
 CC ligands to facilitate drug design and development. The present
 CC sequence is a complementary peptide provided in the specification.
 XX
 PI Roberts GW, Heal JR;
 XX
 DR WPI; 2001-408419/43.
 XX
 PT A set of peptide ligands consisting of specific complementary peptides
 PT to proteins encoded by genes of the human genome, useful in an assay
 PT for screening and identifying of one or more novel peptides which are
 PT drug candidates or pro-drugs -
 XX
 PS Example 6; Page 630; 646pp; English.
 XX
 CC The invention relates to a set of complementary peptide ligands
 CC generated from the human genome. The complementary peptides
 CC interact with their relevant target proteins encoded in the human
 CC genome. They can be used as reagents in drug discovery and as lead
 CC ligands to facilitate drug design and development. The present
 CC sequence is a complementary peptide provided in the specification.
 XX
 Sequence 10 AA;
 XX
 Query Match 91.7%; Score 22; DB 22; Length 10;
 Best Local Similarity 66.7%; Pred. No. 47;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps
 Qy 1 gttxxps 6
 |||
 Db 4 gtssps 9
 XX
 RESULT 6
 AAG97975
 ID AAG97975 standard; Peptide; 10 AA.
 XX
 KW Human; complementary peptide; ligand; drug discovery; drug design.
 AC AAG97975;
 XX
 DR 18-SEP-2001 (first entry)
 XX
 DE Human complementary Peptide, SEQ ID NO: 4170.
 XX
 KW Human complementary peptide; ligand; drug discovery; drug design.
 XX
 OS Homo sapiens.
 XX
 PN WO200142277-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 13-DEC-2000; 2000WO-GB04776.
 XX
 PR 13-DEC-1999; 99GB-0029464.
 XX
 PA (PROT-) PROTEOM LTD.
 XX
 PI Roberts GW, Heal JR;
 XX
 DR WPI; 2001-408419/43.
 XX
 PT A set of peptide ligands consisting of specific complementary peptides
 PT generated from the human genome. The complementary peptides
 PT interact with their relevant target proteins encoded in the human
 CC genome. They can be used as reagents in drug discovery and as lead
 CC ligands to facilitate drug design and development. The present
 CC sequence is a complementary peptide provided in the specification.
 XX
 PS Example 6; Page 631; 646pp; English.
 XX
 DR WPI; 2001-408419/43.
 XX
 PT A set of peptide ligands consisting of specific complementary peptides
 PT to proteins encoded by genes of the human genome, useful in an assay
 PT for screening and identifying of one or more novel peptides which are
 PT drug candidates or pro-drugs -
 XX
 Sequence 10 AA;
 XX
 Query Match 91.7%; Score 22; DB 22; Length 10;
 Best Local Similarity 66.7%; Pred. No. 47;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps
 Qy 1 gttxxps 6

Db 4 gtssps 9

RESULT 7

ID AAR58383 standard; Protein; 38 AA.

XX

AC AAR58383;

XX DT 12-APR-1995 (first entry)

XX DE TSAR binding domain encoded by clone ZnIA7, binds zinc.

XX KW totally synthetic affinity reagent; synthetic; binding domain; effector domain; concatenated heterofunctional protein; linker; direct; rapid; detection; screening; treatment; zinc binding.

XX OS Synthetic.

XX PN WO9418318-A.

XX PD 18-AUG-1994.

XX PF 01-FEB-1994; 94WO-US00977.

XX PR 01-FEB-1993; 93US-0013416.

XX PR 30-DEC-1993; 93US-0176500.

XX PR 31-JAN-1994; 94US-0189331.

XX PA (UYNC-) UNIV NORTH CAROLINA.

XX PI Fowlkes DM, Kay BK;

XX DR WPI; 1994-279739/34.

XX DR N-PSDB; AAQ70472.

XX PT Identifying proteins or peptide(s) which bind a ligand - by screening a recombinant vector library expressing fusion proteins comprising a binding domain and an effector domain

XX PS Claim 38: Page 95; 255pp; English.

XX CC AAR58376-94 show the amino acid sequences of the binding domain of TSAR (Totally Synthetic Affinity Reagents) peptides from the TSAR-9 library. These particular examples bind zinc. The non-variable amino acids at the NH2 and COOH termini are not shown. TSAR peptides are generated using generic oligonucleotides (see AAQ70470-73 for examples). TSARs are concatenated heterofunctional proteins or peptides, comprising at least two functional regions - a binding domain with affinity for a ligand and a second effector peptide portion that is chemically or biologically active. They may further comprise a linker peptide between the 2 domains. The TSARs or consens. comprising a TSAR binding domain can be used in vivo to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or in vivo antibody activity allowing direct and rapid detection in a screening process.

XX Sequence 38 AA;

Query Match 91.7%; Score 22; DB 22; Length 50;

Best Local Similarity 66.7%; Pred. No. 1.6e+02; Mismatches 0; Indels 2; Gaps 0;

Matches 4; Conservative 0; Mismatches 0; Indels 2; Gaps 0;

Qy 1 gtxxps 6

Db 15 gtttsp 25

RESULT 8

RESULT 9

ID AAB53513 standard; protein: 52 AA.

XX AC AAU47468;

XX AC AAB53513;

XX AC AAB53513;

XX DT 09-MAR-2001 (first entry)

DE Human colon cancer antigen protein sequence SEQ ID NO:1053.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX identification; cytostatic; cardioactive; neuroprotective; vulnerable;

XX immunomodulatory; muscular; gynaecological; gastrointestinal;

XX nephrotoxic; antinfective; antibacterial; gene therapy; wound;

XX neural disorder; immune system disorder; muscular disorder;

XX reproductive disorder; gastrointestinal disorder; renal disorder;

XX infectious disease; cardiovascular disorder.

OS Homo sapiens.

XX PN WO20005351-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05883.

XX PR 12-MAR-1999; 990S-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR 2000-587534/55.

XX N-PSDB; AAC98220.

XX PT Colon cancer associated gene sequences, referred to as colon cancer

PT antigens, useful for the treatment, prevention, and diagnosis of colon

PT disorders such as colon cancer -

XX PS Claim 11: Page 1633; 210pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins, called human colon cancer antigens, given in AAB53234 to AAB54006. The human colon cancer antigens can have cytostatic, cardiotactive, muscular, neuroprotective, immunomodulatory, gynaecological, gastrointestinal, pulmonary, nephrotoxic, antinfective and antibacterial activities, and can be used in gene therapy. The colon cancer antigen Polynucleotides, proteins and antibodies to the proteins are useful for the prevention, treatment and diagnosis of colon disorders, such as colon cancer. The polynucleotides may be used in diagnostics and research, such as for chromosome identification, and as hybridisation probes. The proteins may also be used to prevent diseases such as neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, wounds, renal disorders, infectious diseases, and cardiovascular disorders. AAC9764 to AAC98772 and AAB54007 represent sequences used in the exemplification of the present invention.

XX SQ Sequence 52 AA;

Query Match 91.7%; Score 22; DB 21; Length 52;

Best Local Similarity 66.7%; Pred. No. 2.1e+12; Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

Qy 1 gttxps 6

Dy 1 gtssps 6

Db 12 gtatps 17

Query Match 91.7%; Score 22; DB 22; Length 53;

Best Local Similarity 66.7%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gttxps 6

Db 12 gtatps 17

RESULT 11

AAU47468 standard; protein: 53 AA.

ID AAU48691 standard; Protein; 58 AA.
 XX
 AC AAU48691;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #9587.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 WPI: 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 PS Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 9886; 1069pp; English.
 XX
 Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
 XX
 Sequence 58 AA;

AAO10898
 ID AAO10898 standard; Protein; 59 AA.
 XX
 AC AAO10898;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 24790.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematoopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 WPI: 2001-514838/56.
 DR N-PSDB; AAI90829.
 XX
 PS Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 20; SEQ ID NO 24790; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAI179941-AAI193811) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haemopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activation/inhibition activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
 XX
 SQ Sequence 59 AA;

Query Match 1 gtxxps 6
 Best Local Similarity 91.7%; Score 22; DB 22; Length 59;
 Matches 4; Conservative 0; Mismatches 0; Indels 2; Gaps 0; Gaps 0;

Qy 1 gtxxps 6
 Db 49 gtxtps 54

Query Match 1 gtxxps 6
 Best Local Similarity 91.7%; Score 22; DB 22; Length 58;
 Matches 4; Conservative 0; Mismatches 0; Indels 2; Gaps 0; Gaps 0;

Qy 1 gtxxps 6
 Db 6 gtxaps 11

RESULT 13
 ID AAU62078
 XX AC AAU62078 standard; Protein; 61 AA.
 XX
 DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #22974.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteoprotectant; neuroprotectant; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 WPI: 2001-616774/71.
 DR N-PSDB; AAS59623.
 XX
 PS Example 1: SEQ ID No 23273; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 61 AA;
 XX
 Qy Query Match 91.7%; Score 22; DB 22; Length 61;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 42 gtssps 47
 XX
 SQ Sequence 62 AA;
 XX
 Qy Query Match 91.7%; Score 22; DB 22; Length 62;
 Best Local Similarity 66.7%; Pred. No. 2.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 43 gtssps 48
 XX
 RESULT 15
 AAU50474
 ID AAU50474 standard; Protein: 62 AA.
 XX
 AC AAU50474;
 XX
 DT 27-FEB-2002 (first entry)

DT 27-FEB-2002 (first entry)
 XX Propionibacterium acnes immunogenic protein #22875.
 DE XX 24-SEP-1999 (first entry)
 DE T. gondii immunogenic protein.
 KW XX DE T. gondii immunogenic protein.
 KW XX Immunogenic protein; Toxoplasma gondii protein; oocyst shedding; cat;
 KW XX T. gondii infection; enteric apicomplexa oocyst; Cryptosporidium oocyst;
 KW XX Toxoplasma oocyst.
 XX OS XX OS Toxoplasma gondii.
 XX PN XX WO9932633-A1.
 PD XX 01-JUL-1999.
 PF XX 18-DEC-1998; 98WO-US27137.
 XX PR XX 19-DEC-1997; 97US-094825.
 PR XX (HESKA CORP.
 PA XX Lutz SB, Milhausen MJ, Ng RK;
 PA XX WPL; 1999-418930/35.
 PR XX DR N-PSDB; AAX91398.
 PT XX New isolated Toxoplasma gondii nucleic acids used, e.g. to treat
 PT infection caused by this microorganism
 XX Claim 29; Page 319-320; 381pp; English.
 CC XX The invention provides isolated Toxoplasma gondii nucleic acids that
 CC encode immunogenic polypeptides. The T. gondii nucleic acid molecules,
 CC immunogenic proteins and antibodies to the proteins can be used to
 CC prevent T. gondii oocyst shedding in a cat due to infection with
 CC T. gondii. They can be used for preventing T. gondii infection and for
 CC preventing the spread of T. gondii infection. They can also be used for
 CC detecting T. gondii infection. The detection method can be used to detect
 CC parasite cysts or oocysts in feces, e.g. from enteric apicomplexa oocysts.
 CC such as Cryptosporidium oocysts and Toxoplasma oocysts.
 XX Sequence 78 AA;
 SQ XX Query Match 91.7%; Score 22; DB 20; Length 78;
 SQ XX Best Local Similarity 66.7%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;
 SQ XX SO 1 gtxxps 6
 SQ XX Db 50 gtssps 55
 RESULT 17
 ID AAU25543 Standard; Protein; 78 AA.
 XX AC AAU25543;
 XX DT 17-DEC-2001 (first entry)
 DE XX T. gondii immunogenic protein PM2A18-a.
 KW XX Immunogenic protein; oocyst; faeces; enteric apicomplexa oocyst;
 KW XX Cryptosporidium oocyst; Toxoplasma oocyst; Giardia cyst; vaccine;
 KW XX Oocyte Shedding.
 OS XX Toxoplasma gondii.
 XX PN US2001014447-A1.
 XX PD 16-AUG-2001.
 AC AAU25543
 RESULT 16
 ID AAY29072 standard; Protein; 78 AA.
 XX AC AAY29072;

PF 18-DEC-1998; 98US-0216393.
 XX DR 2001-616774/71.
 PR DR N-PSDB; AAS59513.
 XX PT
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 DR PS Example 1; SEQ ID No 2109; 1069pp; English.
 XX
 Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertriosis and osteomyelitis) /uveitis and endophthalmitis
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 78 AA;
 SQ Sequence 80 AA;
 SQ
 Query Match 91.7%; Score 22; DB 22; Length 80;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 50 gtssps 55
 RESULT 18
 AAU40914 AAU40914 standard; Protein; 80 AA.
 ID AAU40914;
 XX AC AAU40914;
 XX DT 13-FEB-2002 (first entry)
 XX DE Propionibacterium acnes immunogenic protein #1810.
 XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertriosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX OS Propionibacterium acnes.
 XX WO200181581-A2.
 XX PD 01-NOV-2001.
 XX PF 20-APR-2001; 2001WO-US12465.
 XX PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 L'maisonneuve J, Zhang Y, Jen S, Carter D;
 PI

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/711.
 DR N-PSDB; AAH35570.
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers -

PT PT
 XX
 PS Claim 11; Page 8369; 9803pp; English.
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytosstatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated P, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent sequences used in the exemplification of the present invention.

CC N.B. Pages 66 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922.

XX SQ Sequence 86 AA;

Query Match 91.7%; Score 22; DB 22; Length 86;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX SQ Sequence 82 AA;

Query Match 91.7%; Score 22; DB 22; Length 82;
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX SQ Sequence 82 AA;

Query Match 91.7%; Score 22; DB 22; Length 86;
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX SQ Sequence 86 AA;

Query Match 91.7%; Score 22; DB 22; Length 86;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX SQ Sequence 86 AA;

RESULT 20
 ID AAG76165 standard; Protein: 86 AA.
 XX AC AAG76165;
 XX DT 03-SEP-2001 (first entry)
 XX DE Human colon cancer antigen protein SEQ ID NO:6929.
 XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 12;
 XX OS Homo sapiens.
 XX PN WO200122920-A2.
 XX PD 05-APR-2001.
 XX PF 28-SEP-2000; 2000WO-US26524.
 XX PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX (HUMA-) HUMAN GENOME SCI. INC.
 XX PR Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX PR WPI; 2001-235357/24.

DR N-PSDB; AAH35570.
 XX PT PT
 XX PS Claim 11; Page 8369; 9803pp; English.
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytosstatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated P, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent sequences used in the exemplification of the present invention.

CC N.B. Pages 66 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922.

XX SQ Sequence 86 AA;

Query Match 91.7%; Score 22; DB 22; Length 86;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX SQ Sequence 82 AA;

Query Match 91.7%; Score 22; DB 22; Length 86;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX SQ Sequence 82 AA;

Query Match 91.7%; Score 22; DB 22; Length 86;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX SQ Sequence 86 AA;

Query Match 91.7%; Score 22; DB 22; Length 86;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX SQ Sequence 86 AA;

RESULT 21
 ID AAM90260 standard; Protein: 90 AA.
 XX AC AAM90260;
 XX DT 07-NOV-2001 (first entry)
 XX DE Human immune/haematopoietic antigen SEQ ID NO:17853.
 XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytosstatic; gene therapy; vaccine; metastasis.
 XX OS Homo sapiens.
 XX PN WO200157182-A2.
 XX PD 09-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US01354.
 XX PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180328.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186550.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-021647.
 PR 07-JUL-2000; 2000US-0216880.

11-JUL-2000; 20000US-0217487.
11-JUL-2000; 20000US-0217496.
14-JUL-2000; 20000US-0218890.
26-JUL-2000; 20000US-0220963.
14-AUG-2000; 20000US-0224518.
14-AUG-2000; 20000US-0224519.
14-AUG-2000; 20000US-0225213.
14-AUG-2000; 20000US-0225214.
14-AUG-2000; 20000US-0225215.
14-AUG-2000; 20000US-0225266.
14-AUG-2000; 20000US-0225267.
14-AUG-2000; 20000US-0225268.
14-AUG-2000; 20000US-0225270.
14-AUG-2000; 20000US-0225271.
14-AUG-2000; 20000US-0225275.
14-AUG-2000; 20000US-0225276.
14-AUG-2000; 20000US-0225277.
14-AUG-2000; 20000US-0225279.
18-AUG-2000; 20000US-0226279.
22-AUG-2000; 20000US-0226681.
22-AUG-2000; 20000US-0226868.
22-AUG-2000; 20000US-0227182.
23-AUG-2000; 20000US-0227009.
30-AUG-2000; 20000US-0228924.
01-SEP-2000; 20000US-0228287.
01-SEP-2000; 20000US-0229443.
01-SEP-2000; 20000US-0229344.
01-SEP-2000; 20000US-0229345.
05-SEP-2000; 20000US-0229509.
05-SEP-2000; 20000US-0229513.
06-SEP-2000; 20000US-0230437.
06-SEP-2000; 20000US-0230438.
08-SEP-2000; 20000US-0231242.
08-SEP-2000; 20000US-0231243.
08-SEP-2000; 20000US-0231244.
08-SEP-2000; 20000US-0231413.
08-SEP-2000; 20000US-0231414.
08-SEP-2000; 20000US-0232080.
12-SEP-2000; 20000US-0231968.
14-SEP-2000; 20000US-0231243.
14-SEP-2000; 20000US-0231244.
14-SEP-2000; 20000US-0231413.
14-SEP-2000; 20000US-0231414.
14-SEP-2000; 20000US-0232239.
14-SEP-2000; 20000US-0232400.
14-SEP-2000; 20000US-0232081.
14-SEP-2000; 20000US-0231063.
14-SEP-2000; 20000US-0231064.
14-SEP-2000; 20000US-0231065.
21-SEP-2000; 20000US-0232223.
21-SEP-2000; 20000US-0232474.
25-SEP-2000; 20000US-0234997.
25-SEP-2000; 20000US-0234998.
26-SEP-2000; 20000US-0235184.
27-SEP-2000; 20000US-0235834.
27-SEP-2000; 20000US-0235836.
29-SEP-2000; 20000US-0236277.
29-SEP-2000; 20000US-0236367.
29-SEP-2000; 20000US-0236368.
29-SEP-2000; 20000US-0236469.
02-OCT-2000; 20000US-0236707.
02-OCT-2000; 20000US-0236902.
02-OCT-2000; 20000US-0237037.
02-OCT-2000; 20000US-0237038.
02-OCT-2000; 20000US-0237039.
13-OCT-2000; 20000US-0239935.
13-OCT-2000; 20000US-0239937.
20-OCT-2000; 20000US-0241096.
20-OCT-2000; 20000US-0241221.
20-OCT-2000; 20000US-0241785.
20-OCT-2000; 20000US-0241786.
20-OCT-2000; 20000US-0241787.
20-OCT-2000; 20000US-0241809.

CC Polynucleotides may be used to produce the secreted (1), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (1) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64/03
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK549/2 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.

XX Sequence 90 AA;

Query Match 91.7%; Score 22; DB 22; Length 91;
 Best Local Similarity 66.7%; pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 1 | | | 6
 Qy 1 gtasps 6
 Db 56 gtssps 61

RESULT 22

ABB39057

ID ABB39057 standard; Peptide; 91 AA.

XX AC ABB39057;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #6563 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PR 30-JAN-2001; 2001WO-US00669.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234667.

XX PR 27-SEP-2000; 2000US-0236319.

XX PR 04-OCT-2000; 2000GB-0024453.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI: 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver

XX Claim 27: SEQ ID NO 31692; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent do not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp://wipo.int/pub/published_pct_sequences.
 XX Sequence 91 AA;

Query Match 91.7%; Score 22; DB 22; Length 91;
 Best Local Similarity 66.7%; pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 56 gtssps 61

RESULT 23

Qy 1 gtxxps 6
 Db 56 gtssps 61

XX AC AAM95131;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen SEQ ID NO: 3789.

XX KW Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PR 17-JAN-2001; 2001WO-US01339.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180528.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0188974.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 11-JUL-2000; 2000US-0217496.

XX PR 14-JUL-2000; 2000US-0218390.

XX PR 26-JUL-2000; 2000US-0220963.

XX PR 26-JUL-2000; 2000US-0220964.

XX PR 14-AUG-2000; 2000US-0224518.

XX PR 14-AUG-2000; 2000US-0224519.

XX PR 14-AUG-2000; 2000US-0225213.

XX PR 14-AUG-2000; 2000US-0225214.

XX PR 14-AUG-2000; 2000US-0225466.

XX PR 14-AUG-2000; 2000US-0225467.

XX PR 14-AUG-2000; 2000US-0225468.

XX PR 14-AUG-2000; 2000US-0225470.

XX PR 14-AUG-2000; 2000US-0225477.

XX PR 14-AUG-2000; 2000US-0225758.

XX PR 14-AUG-2000; 2000US-0225759.

XX PR 18-AUG-2000; 2000US-0226279.

XX PR 22-AUG-2000; 2000US-0226681.

XX PR 22-AUG-2000; 2000US-0226688.

XX PR 22-AUG-2000; 2000US-0227182.

XX PR 23-AUG-2000; 2000US-0227809.

XX PR 30-AUG-2000; 2000US-0228824.

XX PR 01-SEP-2000; 2000US-0229387.

XX PR 01-SEP-2000; 2000US-0229443.

XX PR 01-SEP-2000; 2000US-0229344.

XX PR 05-SEP-2000; 2000US-0229445.

XX PR 05-SEP-2000; 2000US-0229509.

XX PR 05-SEP-2000; 2000US-0229513.

XX PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 20000US-0230438. PR 17-NOV-2000; 20000US-0249244. PR 08-SEP-2000; 20000US-0231242. PR 17-NOV-2000; 20000US-0249245. PR 08-SEP-2000; 20000US-0231243. PR 17-NOV-2000; 20000US-0249264. PR 08-SEP-2000; 20000US-0231244. PR 17-NOV-2000; 20000US-0249265. PR 08-SEP-2000; 20000US-0231413. PR 17-NOV-2000; 20000US-0249297. PR 08-SEP-2000; 20000US-0231414. PR 17-NOV-2000; 20000US-0249299. PR 08-SEP-2000; 20000US-0231080. PR 17-NOV-2000; 20000US-0249300. PR 08-SEP-2000; 20000US-0232081. PR 01-DEC-2000; 20000US-0250160. PR 12-SEP-2000; 20000US-0231968. PR 01-DEC-2000; 20000US-0250391. PR 14-SEP-2000; 20000US-02323397. PR 05-DEC-2000; 20000US-0251030. PR 14-SEP-2000; 20000US-0232398. PR 05-DEC-2000; 20000US-0251988. PR 14-SEP-2000; 20000US-0222399. PR 05-DEC-2000; 20000US-0256719. PR 14-SEP-2000; 20000US-0232400. PR 06-DEC-2000; 20000US-0251479. PR 14-SEP-2000; 20000US-0232401. PR 08-DEC-2000; 20000US-0251856. PR 14-SEP-2000; 20000US-0232402. PR 08-DEC-2000; 20000US-0251868. PR 14-SEP-2000; 20000US-0233064. PR 08-DEC-2000; 20000US-0251869. PR 21-SEP-2000; 20000US-0233165. PR 08-DEC-2000; 20000US-0251989. PR 21-SEP-2000; 20000US-0234123. PR 08-DEC-2000; 20000US-0251990. PR 25-SEP-2000; 20000US-0234274. PR 11-DEC-2000; 20000US-0254097. PR 25-SEP-2000; 20000US-0234397. PR 05-JAN-2001; 20010US-0259678. XX PA (HUMA-) HUMAN GENOME SCI INC. XX PI Rosen CA, Barash SC, Ruben SM; XX DR WPI; 2001-465570/50. XX DR N-PDB; AAL01101. PT Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition -

Claim 11; SEQ ID NO 3789; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a protein of the invention.

Sequence 91 AA;

Query Match 91.7%; Score 22; DB 22; Length 91; Best Local Similarity 66.7%; Pred. No. 3 5e+02; Mismatches 2; Indels 0; Gaps 0;

XX QY 1 gtxxps 6
Db 16 gtxtps 21

RESULT 24
AAM59709
ID AAM59709 standard; Protein; 91 AA.
XX AC AAM59709;
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 31814.
XX KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PP 30-JAN-2001; 2001WO-US00667.
XX PR 04-FEB-2000; 20000US-0180312.

PR 26-MAY-2000; 2000US-02074456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632336.
 PR 21-SEP-2000; 2000US-0234677.
 PR 27-SEP-2000; 2000US-0236339.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR; 2001-483446/52.
 XX PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX SQ Example 4; SEQ ID NO: 31814; 650pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX SQ Sequence 91 AA;
 XX Query Match 91.7%; Score 22; DB 22;
 XX Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 XX Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX Qy 1 9ttxps 6
 XX Db 56 9tssps 61
 XX RESULT 26
 ID AAM32548 standard; Protein; 91 AA.
 XX ID AAM32548 standard; Protein; 91 AA.
 XX AC AAM32548;
 XX DT 17-OCT-2001 (first entry)
 XX DE Peptide #6585 encoded by probe for measuring placental gene expression.
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX KW genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200157272-A2.
 XX PD 09-AUG-2001.
 XX PP 30-JAN-2001; 2001WO-US00663.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX DE 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632336.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000US-0236359.
 XX PR 09-AUG-2001.
 XX PR 30-JAN-2001; 2001WO-US00668.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632336.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR; 2001-4888900/53.
 XX SQ Sequence 91 AA;

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX PS Example 4; SEQ ID NO: 32595; 658pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX SQ Sequence 91 AA;

Query Match 91.7%; Score 22; DB 22;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 9ttxps 6
 Db 56 9tssps 61

RESULT 26
 ID AAM32548 standard; Protein; 91 AA.

XX AC AAM32548;
 XX DT 17-OCT-2001 (first entry)
 XX DE Peptide #6585 encoded by probe for measuring placental gene expression.
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX KW genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200157272-A2.
 XX PD 09-AUG-2001.
 XX PP 30-JAN-2001; 2001WO-US00663.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX DE 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632336.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR; 2001-4888900/53.
 XX SQ Sequence 91 AA;

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX PS Claim 27; SEQ ID NO 32817; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENPs;
 CC see AAI31315-AAI57346). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX SQ Sequence 91 AA;

Query Match 91.7%; Score 22; DB 22; Length 91;
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 56 gtssps 61

RESULT 27

ID AAM90809 standard; Protein; 106 AA.
 XX
 AC AAM90809;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen SEQ ID NO:18402.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0189123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0221680.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0224519.
 PR 22-AUG-2000; 2000US-0225270.
 PR 22-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227109.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229387.
 PR 01-SEP-2000; 2000US-0229434.
 PR 01-SEP-2000; 2000US-0229445.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-02230437.
 PR 08-SEP-2000; 2000US-02231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0233081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234220.
 PR 21-SEP-2000; 2000US-0234274.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0244998.
 PR 26-SEP-2000; 2000US-025484.
 PR 27-SEP-2000; 2000US-0255834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0244647.
 PR 08-NOV-2000; 2000US-0244747.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0246527.
 PR 17-NOV-2000; 2000US-0246532.
 PR 17-NOV-2000; 2000US-0246609.
 PR 17-NOV-2000; 2000US-0246610.
 PR 17-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0246617.
 PR 17-NOV-2000; 2000US-0246620.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.

PR	17-NOV-2000;	2000US-0249218.	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 28293.
PR	17-NOV-2000;	2000US-024924.	XX	Protein identification; signal transduction pathway; metabolic pathway;
PR	17-NOV-2000;	2000US-0249245.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	17-NOV-2000;	2000US-0249264.	KW	termination sequence.
PR	17-NOV-2000;	2000US-0249265.	KW	
PR	17-NOV-2000;	2000US-0249265.	XX	
PR	17-NOV-2000;	2000US-0249297.	XX	Arabidopsis thaliana.
PR	17-NOV-2000;	2000US-0249309.	OS	
PR	01-DEC-2000;	2000US-0250160.	XX	
PR	01-DEC-2000;	2000US-02501391.	PN	EP1033405-A2.
PR	05-DEC-2000;	2000US-0251030.	XX	
PR	05-DEC-2000;	2000US-025198.	PD	06-SEP-2000.
PR	05-DEC-2000;	2000US-025198.	XX	25-FEB-2000;
PR	05-DEC-2000;	2000US-0256719.	PF	2000EP-0301439.
PR	08-DEC-2000;	2000US-0251479.	XX	
PR	08-DEC-2000;	2000US-0251856.	PR	25-FEB-1999;
PR	08-DEC-2000;	2000US-0251868.	PR	05-MAR-1999;
PR	08-DEC-2000;	2000US-0251879.	PR	09-MAR-1999;
PR	08-DEC-2000;	2000US-025198.	PR	23-MAR-1999;
PR	08-DEC-2000;	2000US-0251990.	PR	25-MAR-1999;
PR	11-DEC-2000;	2000US-0251990.	PR	29-MAR-1999;
PR	05-JAN-2001;	2000US-0254197.	PR	01-APR-1999;
PR	05-JAN-2001;	2000US-0259678.	PR	01-APR-1999;
XX	(HUMA -) HUMAN GENOME SCI INC.	XX	99US-0128334.	
PA		PA	99US-0128714.	
XX		PR	08-APR-1999;	
PI	Rosen CA, Barash SC, Ruben SM;	PR	16-APR-1999;	
XX		PR	19-APR-1999;	
DR	WPI; 2001-483426/52.	PR	21-APR-1999;	
DR	N-PSDB: AAK63590.	PR	23-APR-1999;	
XX		PR	01-APR-1999;	
PR	Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -	PR	06-APR-1999;	
PR		PR	08-APR-1999;	
PR		PR	16-APR-1999;	
PR		PR	19-APR-1999;	
PR		PR	23-APR-1999;	
PR		PR	04-MAY-1999;	
PR		PR	05-MAY-1999;	
PS	Claim 11; SEQ ID NO 18402; 3071pp + Sequence Listing; English.	PR	06-MAY-1999;	
XX		PR	06-MAY-1999;	
CC	AAK5951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM62170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAMB2169 represent sequences used in the exemplification of the present invention.	XX	99US-0132486.	
CC		XX	99US-0132487.	
CC		CC	99US-0132483.	
CC		CC	99US-013256.	
CC		CC	99US-0134218.	
CC		CC	99US-013419.	
CC		CC	99US-0134221.	
CC		CC	99US-0134370.	
CC		CC	99US-0134768.	
CC		CC	99US-0134941.	
CC		CC	99US-0135124.	
CC		CC	99US-0135353.	
CC		CC	99US-0135629.	
CC		CC	99US-0136021.	
CC		CC	99US-0136392.	
CC		CC	99US-0136782.	
CC		CC	99US-0137722.	
CC		CC	99US-0137728.	
CC		CC	99US-0137724.	
CC		CC	99US-0138094.	
CC		CC	99US-013840.	
CC		CC	99US-013919.	
CC		CC	99US-0139452.	
CC		CC	99US-0139453.	
CC		CC	99US-0139454.	
CC		CC	99US-0139455.	
CC		CC	99US-0139456.	
CC		CC	99US-0139457.	
CC		CC	99US-0139458.	
CC		CC	99US-0139459.	
CC		CC	99US-0139460.	
CC		CC	99US-0139461.	
CC		CC	99US-0139462.	
CC		CC	99US-0139463.	
CC		CC	99US-0139465.	
CC		CC	99US-0139467.	
CC		CC	99US-0139468.	
CC		CC	99US-0139469.	
CC		CC	99US-0139470.	
CC		CC	99US-0139471.	
CC		CC	99US-0139472.	
CC		CC	99US-0139473.	
CC		CC	99US-0139474.	
CC		CC	99US-0139475.	
CC		CC	99US-0139476.	
CC		CC	99US-0139477.	
CC		CC	99US-0139478.	
CC		CC	99US-0139479.	
CC		CC	99US-0139480.	
CC		CC	99US-0139481.	
CC		CC	99US-0139482.	
CC		CC	99US-0139483.	
CC		CC	99US-0139484.	
CC		CC	99US-0139485.	
CC		CC	99US-0139486.	
CC		CC	99US-0139487.	
CC		CC	99US-0139488.	
CC		CC	99US-0139489.	
CC		CC	99US-0139490.	
CC		CC	99US-0139491.	
CC		CC	99US-0139492.	
CC		CC	99US-0139493.	
CC		CC	99US-0139494.	
CC		CC	99US-0139495.	
CC		CC	99US-0139496.	
CC		CC	99US-0139497.	
CC		CC	99US-0139498.	
CC		CC	99US-0139499.	
CC		CC	99US-0139500.	
CC		CC	99US-0139501.	
CC		CC	99US-0139502.	
CC		CC	99US-0139503.	
CC		CC	99US-0139504.	
CC		CC	99US-0139505.	
CC		CC	99US-0139506.	
CC		CC	99US-0139507.	
CC		CC	99US-0139508.	
CC		CC	99US-0139509.	
CC		CC	99US-0139510.	
CC		CC	99US-0139511.	
CC		CC	99US-0139512.	
CC		CC	99US-0139513.	
CC		CC	99US-0139514.	
CC		CC	99US-0139515.	
CC		CC	99US-0139516.	
CC		CC	99US-0139517.	
CC		CC	99US-0139518.	
CC		CC	99US-0139519.	
CC		CC	99US-0139520.	
CC		CC	99US-0139521.	
CC		CC	99US-0139522.	
CC		CC	99US-0139523.	
CC		CC	99US-0139524.	
CC		CC	99US-0139525.	
CC		CC	99US-0139526.	
CC		CC	99US-0139527.	
CC		CC	99US-0139528.	
CC		CC	99US-0139529.	
CC		CC	99US-0139530.	
CC		CC	99US-0139531.	
CC		CC	99US-0139532.	
CC		CC	99US-0139533.	
CC		CC	99US-0139534.	
CC		CC	99US-0139535.	
CC		CC	99US-0139536.	
CC		CC	99US-0139537.	
CC		CC	99US-0139538.	
CC		CC	99US-0139539.	
CC		CC	99US-0139540.	
CC		CC	99US-0139541.	
CC		CC	99US-0139542.	
CC		CC	99US-0139543.	
CC		CC	99US-0139544.	
CC		CC	99US-0139545.	
CC		CC	99US-0139546.	
CC		CC	99US-0139547.	
CC		CC	99US-0139548.	
CC		CC	99US-0139549.	
CC		CC	99US-0139550.	
CC		CC	99US-0139551.	
CC		CC	99US-0139552.	
CC		CC	99US-0139553.	
CC		CC	99US-0139554.	
CC		CC	99US-0139555.	
CC		CC	99US-0139556.	
CC		CC	99US-0139557.	
CC		CC	99US-0139558.	
CC		CC	99US-0139559.	
CC		CC	99US-0139560.	
CC		CC	99US-0139561.	
CC		CC	99US-0139562.	
CC		CC	99US-0139563.	
CC		CC	99US-0139564.	
CC		CC	99US-0139565.	
CC		CC	99US-0139566.	
CC		CC	99US-0139567.	
CC		CC	99US-0139568.	
CC		CC	99US-0139569.	
CC		CC	99US-0139570.	
CC		CC	99US-0139571.	
CC		CC	99US-0139572.	
CC		CC	99US-0139573.	
CC		CC	99US-0139574.	
CC		CC	99US-0139575.	
CC		CC	99US-0139576.	
CC		CC	99US-0139577.	
CC		CC	99US-0139578.	
CC		CC	99US-0139579.	
CC		CC	99US-0139580.	
CC		CC	99US-0139581.	
CC		CC	99US-0139582.	
CC		CC	99US-0139583.	
CC		CC	99US-0139584.	
CC		CC	99US-0139585.	
CC		CC	99US-0139586.	
CC		CC	99US-0139587.	
CC		CC	99US-0139588.	
CC		CC	99US-0139589.	
CC		CC	99US-0139590.	
CC		CC	99US-0139591.	
CC		CC	99US-0139592.	
CC		CC	99US-0139593.	
CC		CC	99US-0139594.	
CC		CC	99US-0139595.	
CC		CC	99US-0139596.	
CC		CC	99US-0139597.	
CC		CC	99US-0139598.	
CC		CC	99US-0139599.	
CC		CC	99US-0139600.	
CC		CC	99US-0139601.	
CC		CC	99US-0139602.	
CC		CC	99US-0139603.	
CC		CC	99US-0139604.	
CC		CC	99US-0139605.	
CC		CC	99US-0139606.	
CC		CC	99US-0139607.	
CC		CC	99US-0139608.	
CC		CC	99US-0139609.	
CC		CC	99US-0139610.	
CC		CC	99US-0139611.	
CC		CC	99US-0139612.	
CC		CC	99US-0139613.	
CC		CC	99US-0139614.	
CC		CC	99US-0139615.	
CC		CC	99US-0139616.	
CC		CC	99US-0139617.	
CC		CC	99US-0139618.	
CC		CC	99US-0139619.	
CC		CC	99US-0139620.	
CC		CC	99US-0139621.	
CC		CC	99US-0139622.	
CC		CC	99US-0139623.	
CC		CC	99US-0139624.	
CC		CC	99US-0139625.	
CC		CC	99US-0139626.	
CC		CC	99US-0139627.	
CC		CC	99US-0139628.	
CC		CC	99US-0139629.	
CC		CC	99US-0139630.	
CC		CC	99US-0139631.	
CC		CC	99US-0139632.	
CC		CC	99US-0139633.	
CC		CC	99US-0139634.	
CC		CC	99US-0139635.	
CC		CC	99US-0139636.	
CC		CC	99US-0139637.	
CC		CC	99US-0139638.	
CC		CC	99US-0139639.	
CC		CC	99US-0139640.	
CC		CC	99US-0139641.	
CC		CC	99US-0139642.	
CC		CC	99US-0139643.	
CC		CC	99US-0139644.	
CC		CC	99US-0139645.	
CC		CC	99US-0139646.	
CC		CC	99US-0139647.	
CC		CC	99US-0139648.	
CC		CC	99US-0139649.	
CC		CC	99US-0139650.	
CC		CC	99US-0139651.	
CC		CC	99US-0139652.	
CC		CC	99US-0139653.	
CC		CC	99US-0139654.	
CC		CC	99US-0139655.	
CC		CC	99US-0139656.	
CC		CC	99US-0139657.	
CC		CC	99US-0139658.	
CC		CC	99US-0139659.	
CC		CC	99US-0139660.	
CC		CC	99US-0139661.	
CC		CC	99US-0139662.	
CC		CC	99US-0139663.	
CC		CC	99US-0139664.	
CC		CC	99US-0139665.	
CC		CC	99US-0139666.	
CC		CC	99US-0139667.	
CC		CC	99US-0139668.	
CC		CC	99US-0139669.	
CC		CC	99US-0139670.</	

PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153158.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142354.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-014325.	PR	04-OCT-1999;	99US-0157117.
PR	13-JUL-1999;	99US-0143542.	PR	05-OCT-1999;	99US-0157753.
PR	14-JUL-1999;	99US-0144624.	PR	06-OCT-1999;	99US-0157865.
PR	15-JUL-1999;	99US-0144705.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144805.	PR	08-OCT-1999;	99US-0158232.
PR	16-JUL-1999;	99US-0144899.	PR	12-OCT-1999;	99US-0158469.
PR	19-JUL-1999;	99US-0142920.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144086.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144325.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159296.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159297.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159298.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159337.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159638.
PR	20-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0144884.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145016.	PR	22-OCT-1999;	99US-0160767.
PR	21-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0144335.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0144336.	PR	21-OCT-1999;	99US-0160815.
PR	22-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	23-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145309.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145313.	PR	26-OCT-1999;	99US-0161359.
PR	27-JUL-1999;	99US-0145318.	PR	26-OCT-1999;	99US-0161360.
PR	28-JUL-1999;	99US-0145386.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0145388.	PR	28-OCT-1999;	99US-0161992.
PR	02-AUG-1999;	99US-0145389.	PR	28-OCT-1999;	99US-0161993.
PR	03-AUG-1999;	99US-0147038.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-014704.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147102.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147260.	RESULT	29	
PR	06-AUG-1999;	99US-0147303.	Qy	1	9ttxps 6
PR	06-AUG-1999;	99US-0147416.	Db	71	9ttxps 76
PR	09-AUG-1999;	99US-0147493.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-014884.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149930.			
PR	23-AUG-1999;	99US-0150566.			
PR	25-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			

Query Match Score 91.7%;
Best Local Similarity 66.7%;
Pred. No. 4.1e-02;
Matches 4; Conservative 0;
Mismatches 2; Indels 0;
Gaps 0;

Query Match Score 22; DB 21; Length 109;
Best Local Similarity 66.7%;
Pred. No. 4.1e-02;

Qy 1 9ttxps 6
Db 71 9ttxps 76

RESULT 29
AAM92684
ID AAM92684 standard; Protein; 110 AA.
XX
AC AAM92684;
XX
DE 06-NOV-2001 (first entry)
AAM92684
ID AAM92684 standard; Protein; 110 AA.
XX
AC AAM92684;
XX
DE Human digestive system antigen SEQ ID NO : 2033 .
AAM92684
ID AAM92684 standard; Protein; 110 AA.
XX
AC AAM92684;
XX
DE Human digestive system antigen; gene therapy; cancer; appendicitis;
ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
digestive system disorder; Meckel's diverticulum.
XX
OS Homo sapiens.

XX	WO200155314-A2.	27-SEP-2000;	20000US-0235836.		
PN		29-SEP-2000;	20000US-0236227.		
XX		29-SEP-2000;	20000US-0233367.		
PD	02-AUG-2001.	29-SEP-2000;	20000US-0233368.		
XX		29-SEP-2000;	20000US-0233369.		
PF	17-JAN-2001;	2001IW0-US01324.	PR	29-SEP-2000;	20000US-0236200.
XX		02-OCT-2000;	20000US-023670.		
PR	31-JAN-2000;	20000US-0179065.	PR	02-OCT-2000;	20000US-0236802.
PR	04-FEB-2000;	20000US-0184664.	PR	02-OCT-2000;	20000US-0237038.
PR	02-MAR-2000;	20000US-0186350.	PR	02-OCT-2000;	20000US-0237039.
PR	16-MAR-2000;	20000US-018984.	PR	02-OCT-2000;	20000US-0237040.
PR	17-MAR-2000;	20000US-0190077.	PR	13-OCT-2000;	20000US-0239935.
PR	18-APR-2000;	20000US-0198123.	PR	13-OCT-2000;	20000US-0239937.
PR	19-MAY-2000;	20000US-0205515.	PR	20-OCT-2000;	20000US-0241809.
PR	07-JUN-2000;	20000US-0209467.	PR	20-OCT-2000;	20000US-0241821.
PR	28-JUN-2000;	20000US-021488.	PR	20-OCT-2000;	20000US-0241786.
PR	30-JUN-2000;	20000US-0215135.	PR	20-OCT-2000;	20000US-0241787.
PR	07-JUL-2000;	20000US-0216647.	PR	20-OCT-2000;	20000US-0241808.
PR	07-JUL-2000;	20000US-0216880.	PR	20-OCT-2000;	20000US-0241809.
PR	11-JUL-2000;	20000US-0217487.	PR	20-OCT-2000;	20000US-0241826.
PR	11-JUL-2000;	20000US-0217496.	PR	01-NOV-2000;	20000US-0244617.
PR	14-JUL-2000;	20000US-0218290.	PR	08-NOV-2000;	20000US-0244674.
PR	26-JUL-2000;	20000US-0220963.	PR	08-NOV-2000;	20000US-0244675.
PR	26-JUL-2000;	20000US-0220964.	PR	08-NOV-2000;	20000US-0244676.
PR	14-AUG-2000;	20000US-0224518.	PR	08-NOV-2000;	20000US-0244677.
PR	14-AUG-2000;	20000US-0224519.	PR	08-NOV-2000;	20000US-0244678.
PR	14-AUG-2000;	20000US-0225213.	PR	08-NOV-2000;	20000US-0246523.
PR	14-AUG-2000;	20000US-0225214.	PR	08-NOV-2000;	20000US-0246524.
PR	14-AUG-2000;	20000US-022526.	PR	08-NOV-2000;	20000US-0246525.
PR	14-AUG-2000;	20000US-0225267.	PR	08-NOV-2000;	20000US-0246526.
PR	14-AUG-2000;	20000US-0225468.	PR	08-NOV-2000;	20000US-0246527.
PR	14-AUG-2000;	20000US-0225477.	PR	08-NOV-2000;	20000US-0246528.
PR	14-AUG-2000;	20000US-0225757.	PR	08-NOV-2000;	20000US-0245532.
PR	14-AUG-2000;	20000US-0225758.	PR	08-NOV-2000;	20000US-0246609.
PR	14-AUG-2000;	20000US-0225759.	PR	08-NOV-2000;	20000US-0246510.
PR	18-AUG-2000;	20000US-0226279.	PR	08-NOV-2000;	20000US-0246611.
PR	22-AUG-2000;	20000US-0226681.	PR	08-NOV-2000;	20000US-0246613.
PR	22-AUG-2000;	20000US-0226868.	PR	17-NOV-2000;	20000US-0249208.
PR	23-AUG-2000;	20000US-0227182.	PR	17-NOV-2000;	20000US-0249209.
PR	01-SEP-2000;	20000US-0228924.	PR	17-NOV-2000;	20000US-0249210.
PR	01-SEP-2000;	20000US-0229287.	PR	17-NOV-2000;	20000US-0249211.
PR	01-SEP-2000;	20000US-0229343.	PR	17-NOV-2000;	20000US-0249212.
PR	01-SEP-2000;	20000US-0229344.	PR	17-NOV-2000;	20000US-0249213.
PR	01-SEP-2000;	20000US-0229345.	PR	17-NOV-2000;	20000US-0249214.
PR	05-SEP-2000;	20000US-0229509.	PR	17-NOV-2000;	20000US-0249215.
PR	05-SEP-2000;	20000US-0229513.	PR	17-NOV-2000;	20000US-0249216.
PR	06-SEP-2000;	20000US-0230437.	PR	17-NOV-2000;	20000US-0249217.
PR	06-SEP-2000;	20000US-0230438.	PR	17-NOV-2000;	20000US-0249218.
PR	08-SEP-2000;	20000US-0231242.	PR	17-NOV-2000;	20000US-024924.
PR	08-SEP-2000;	20000US-0231243.	PR	17-NOV-2000;	20000US-0249245.
PR	08-SEP-2000;	20000US-0231244.	PR	17-NOV-2000;	20000US-0249264.
PR	14-SEP-2000;	20000US-0231249.	PR	01-DEC-2000;	20000US-0250391.
PR	14-SEP-2000;	20000US-0231249.	PR	05-DEC-2000;	20000US-0251030.
PR	14-SEP-2000;	20000US-0232398.	PR	05-DEC-2000;	20000US-0251388.
PR	14-SEP-2000;	20000US-0232399.	PR	05-DEC-2000;	20000US-0256719.
PR	14-SEP-2000;	20000US-0234240.	PR	06-DEC-2000;	20000US-0251479.
PR	14-SEP-2000;	20000US-0234241.	PR	08-DEC-2000;	20000US-025156.
PR	14-SEP-2000;	20000US-0233063.	PR	08-DEC-2000;	20000US-0251868.
PR	14-SEP-2000;	20000US-0233064.	PR	08-DEC-2000;	20000US-0251869.
PR	14-SEP-2000;	20000US-0233065.	PR	08-DEC-2000;	20000US-0251989.
PR	21-SEP-2000;	20000US-0234243.	PR	08-DEC-2000;	20000US-0251990.
PR	21-SEP-2000;	20000US-0234274.	PR	11-DEC-2000;	20000US-0254097.
PR	25-SEP-2000;	20000US-0234397.	PR	05-JAN-2001;	2001US-0259678.
PR	25-SEP-2000;	20000US-0234398.	XX		
PR	26-SEP-2000;	20000US-023584.	PA		
PR	27-SEP-2000;	20000US-0235834.	XX		

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-50263/055.
 DR N-FSDB; AAK88457.
 XX

XX Polynucleotides encoding digestive system antigens, useful for diagnosing, treating, preventing and/or prognosing disorders of the digestive system, particularly cancer and cancer metastases -
 XX

PS Claim 11; SEQ ID NO 2033; 986pp; English.
 XX

CC The present invention provides the protein and coding sequences of a number of human digestive system antigens. These can be used in the diagnosis, treatment and prevention of digestive system disorders, including cancer, Meckel's diverticulum, bacterial or parasitic infections, appendicitis, Hirschsprung's disease, chronic colitis or ulcerative colitis. The present sequence is a digestive system antigen of the invention.
 XX

Sequence 110 AA;

Query Match 91.7%; Score 22; DB 22; Length 110;
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 13 gtssps 18

RESULT 30
 ID AAY12647 standard; Protein; 116 AA.
 XX
 AC AAY12647;
 XX DT 22-JUN-1999 (first entry)
 XX

DE Human 5' EST secreted protein SEQ ID NO: 312 from WO 9906553.
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haemopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; antiinflammatory; tumour inhibition; antitumour.
 XX

OS Homo sapiens.
 PN WO9906553-A2.
 XX
 PD 11-FEB-1999.
 XX PF 31-JUL-1998; 98WO-IB01237.
 XX PR 01-AUG-1997; 97US-0905051.
 XX WPI; 1999-153783/13.
 PA (GEST) GENSET.
 XX
 PI Duclert A, Dumas Milne Edwards J, Lacroix B;
 XX DR N-PSDB; AAX41505.
 XX

XX New nucleic acids encoding human secreted proteins - obtained from cDNA libraries derived from umbilical cord, lymph ganglia, lymphocytes and placental tissue
 XX

PS Claim 34; Page 397-398; 411pp; English.
 XX AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for human secreted proteins, and encode the proteins given in AAY12521 to

CC AAY12668, respectively. The proteins given represent the signal peptide and an N-terminal fragment of a secreted protein. The nucleic acid sequences can be used for producing secreted human gene products. They can also be used to develop products for diagnosis and therapy. The proteins obtained may have cytosolic activity, cell proliferation/differentiation activity, haemopoiesis regulating activity, tissue growth regulating activity, reproductive hormone regulating activity, chemotactic/ chemokinetic activity, haemostatic and thrombolytic activity, receptor/ ligand activity, antiinflammatory activity, tumour inhibition activity or other activities. The products can be used in forensic, gene therapy and chromosome mapping procedures. The sequences can also be used for obtaining corresponding promoter sequences. The nucleic acids encoding the signal peptide can be used for directing extracellular secretion of a polypeptide or the insertion of a polypeptide into a membrane, or importing a polypeptide into a cell.
 XX

SQ Sequence 116 AA;

Query Match 91.7%; Score 22; DB 20; Length 116;
 Best Local Similarity 66.7%; Pred. No. 4.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 53 gtssps 58

RESULT 31
 ID AAU65364 standard; Protein; 118 AA.
 XX
 AC AAU65364;
 XX DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #26260.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; dermatological; osteopathic; enzyme linked immunosorbent assay;
 XX
 KW Propionibacterium acnes.
 XX OS
 PN WO2001181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-1999047P.
 PR 02-JUN-2000; 2000US-209841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORY-) CORIXA CORP.
 XX
 PI Skeik YAW, Persing DH, Mitcham JL, Wang SS, Bhattacharya A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX DR WPI; 2001-616774/71.
 XX DR N-PSDB; AAS55666.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for vaccinating against and diagnosing infections, especially useful for treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 26559; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU8017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 118 AA;

Query Match 91.7%; Score 22; DB 22; Length 118;
 Best Local Similarity 66.7%; Pred. No. 4.e-02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 qtxxps 6
 Db 111 gtatps 116

RESULT 32
 AAU50960 standard; Protein: 122 AA.
 XX AAU50960;
 AC
 XX DT 27-FEB-2002 (first entry)
 XX DE Propionibacterium acnes immunogenic protein #11856.
 XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX OS Propionibacterium acnes.
 XX PN WO200181581-A2.
 XX PD 01-NOV-2001.
 XX PF 20-APR-2001; 2001WO-US12865.
 XX PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX PA (CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI I'maisonneue J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/71.
 DR N-PSDB; AAS59549.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX Example 1; SEQ ID No 12155; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 122 AA;

Query Match 91.7%; Score 22; DB 22; Length 122;
 Best Local Similarity 66.7%; Pred. No. 4.6e-02; Indels 2; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 qtxxps 6
 Db 118 qtxxps 23

RESULT 33
 AAB32475 standard; Protein: 157 AA.
 ID AAB32476
 XX AC AAB32476;
 AC XX DT 26-JUN-2001 (first entry)
 XX DE Human protein sequence SEQ ID NO:10552.
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-030253.
 PR 11-JUN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX PS Claim 8; SEQ ID 10552; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of CC the 5602 nucleotide sequences defined in the specification; or where the CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination CC of an oligonucleotide comprising a sequence complementary to the CC complementary strand of a polynucleotide which comprises a 5'-end CC sequence and an oligonucleotide comprising a sequence complementary to a CC polynucleotide which comprises a 3'-end sequence, where the CC oligonucleotide comprises at least 15 nucleotides and the combination of CC the 5'-end sequence/3'-end sequence is selected from those defined in CC the specification. The primer sets can be used in antisense therapy and CC in gene therapy. The primers are useful for synthesising polynucleotides, CC particularly full-length cDNAs. The primers are also useful for the CC detection and/or diagnosis of the abnormality of the proteins encoded by CC the full-length cDNAs. The primers allow obtaining of the full-length CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 CC represent oligonucleotides, all of which are used in the exemplification CC of the present invention.

XX Sequence 157 AA;

Query Match 91.7%; Score 22; DB 22; Length 157;
Best Local Similarity 66.7%; Pred. No. 5.7e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2;

Qy 1 gtxxps 6
Db 76 gtaaps 81

RESULT 34
ID AAB32650 standard; Protein; 160 AA.
XX AC AAB32650;
XX DT 25-JAN-2001 (first entry)
XX DE Eucalyptus grandis transcription factor protein sequence #108.
XX KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeobox; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;
KW type 2 Cys2Hs2; CCAAT box element; MYB.
XX OS Eucalyptus grandis.
XX PN WO20053724-A2.
XX PD 14-SEP-2000.
XX PF 09-MAR-2000; 2000WO-US06112.
XX PR 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.

XX PR WPI; 2000-579369/54.

XX PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
PT -
XX PI Wood M, McGrath A, Shenk MA, Glenn M;
XX DR WPI; 2000-579369/54.

XX XX PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
PT -
XX PS Claim 8; Page 243; 747pp; English.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.

CC The present invention relates to novel plant transcription factors from CC Eucalyptus grandis or Pinus radiata. The present sequence is one such CC transcription factor. The transcription factor may be used to produce a CC plant having modified gene expression such as a woody plant e.g. a CC eucalyptus, pine, acacia, sweetgum, teak, or mahogany species or CC to modify the activity of a polypeptide in a plant. The transcription CC factors of the present invention are members from the following families CC of regulatory proteins: bZIP, bZIP family of C-box binding factors, basic CC helix-loop-helix zipper, homeotic/homeobox/MADS, homeodomain CC zipper, LIM domain, AP2 and EREBS, zinc finger elements and MYB.
CC Cys2Hs2, CCAAT box elements and MYB.
CC Sequence 160 AA;
SQ 1 gtxxps 6
Db 112 gtaaps 117

Query Match 91.7%; Score 22; DB 21; Length 160;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 2;
Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 112 gtaaps 117

RESULT 35
ID AAG24568 standard; Protein; 160 AA.
XX AC AAG24568;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28292.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 11-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132456.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134914.
 PR 20-MAY-1999; 99US-0135154.
 PR 21-MAY-1999; 99US-0135333.
 PR 24-MAY-1999; 99US-0135639.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136322.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137562.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138034.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139465.
 PR 18-JUN-1999; 99US-0139467.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139839.
 PR 23-JUN-1999; 99US-0140533.
 PR 24-JUN-1999; 99US-0140634.
 PR 24-JUN-1999; 99US-0140635.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143342.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144086.
 PR 21-JUL-1999; 99US-0144335.
 PR 22-JUL-1999; 99US-0144352.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 20-JUL-1999; 99US-0144984.
 PR 21-JUL-1999; 99US-0144812.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-014793.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148111.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148365.
 PR 16-AUG-1999; 99US-0148384.
 PR 17-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149375.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149829.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150366.
 PR 26-AUG-1999; 99US-0150884.
 PR 26-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-015103.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152463.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155309.
 PR 24-SEP-1999; 99US-0155486.
 PR 28-SEP-1999; 99US-0155659.
 PR 29-SEP-1999; 99US-0156458.
 PR 04-OCT-1999; 99US-0156596.
 PR 05-OCT-1999; 99US-0157117.
 PR 06-OCT-1999; 99US-0157753.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 14-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159299.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159337.
 PR 21-OCT-1999; 99US-0159338.
 PR 21-OCT-1999; 99US-0159345.
 PR 22-OCT-1999; 99US-0160741.
 PR 22-OCT-1999; 99US-0160747.
 PR 22-OCT-1999; 99US-0160768.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 91.7%; Score 22; DB 21; Length 160;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; OS Synthetic.
 Qy 1 gtxxps 6
 Db 122 gtttts 127

RESULT 36
 ID AAW32476 standard; Protein: 166 AA.
 XX
 AC AAW32476;
 XX
 DT 15-JAN-1998 (first entry)
 XX
 DE BBC6 protein for regulating cell death.
 XX
 KW BBC6 gene; cell death; cell cycle; Bc12; human.
 XX
 OS Homo sapiens.
 XX
 PN US5663316-A.
 XX
 PD 02-SEP-1997.
 XX
 PF 18-JUN-1996; 96US-0665617.
 XX
 PR 18-JUN-1996; 96US-0665617.
 XX
 PA (CLON-) CLONTECH LAB INC.
 XX
 PI Xudong Y;
 XX
 WPI: 1997-447980/41.
 DR N-PSDB; AAT91561.
 PS Claim 1; Column 11-12; 7pp; English.
 XX
 PT Isolated BBC6 gene - encodes a protein that regulates cell death
 PT through interaction with Bc1-2
 XX
 CC The present sequence represents a protein of 166 amino acids. The
 sequence is disclosed as being a protein called BBC6 which regulates
 cell death through interaction with Bc1-2. The DNA may be used for the
 production of the recombinant protein, which can be used in unspecified
 therapeutic or diagnostic procedures, as a molecular weight marker, and
 to raise antibodies that can be used in unspecified diagnostic or
 therapeutic applications and to reduce or eliminate the biological
 activity of the BBC6 protein in vivo.
 XX
 SQ Sequence 166 AA;

Query Match 91.7%; Score 22; DB 17; Length 168;
 Best Local Similarity 66.7%; Pred. No. 6.e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 160 gtaaps 165

RESULT 38
 AAU38749

AAR95975
 ID AAR95975 standard; peptide; 168 AA.
 XX
 AC AAR95975;
 XX
 DT 19-FEB-1997 (first entry)
 XX
 DE Nucleic acid recognition unit #11.
 XX
 KW Target binding assembly; nucleic acid recognition unit; NF-kappa-B; SPI;
 KW TATA; human papillomavirus; HPV; HPV E2; human immunodeficiency virus;
 KW HIV LTR; Tat binding unit; probe; assembly sequence; asymmetry sequence;
 KW nuclear localisation signal sequence; human cell; HPV LTR; HIV; bacteria;
 KW virus.
 XX
 OS Synthetic.
 XX
 PN WO9517956-A2.
 XX
 PD 13-JUN-1996.
 XX
 PF 07-DEC-1995; 95WO-US15944.
 XX
 PR 09-DEC-1994; 94US-0353476.
 XX
 PA (GENE-) GENE POOL INC.
 XX
 PI Weininger AM, Weininger S;
 XX
 DR WPI; 1996-287199/29.
 XX
 PT Probe nucleic acids, target binding assemblies, etc - for detection
 and localisation of specific nucleic acid sequences, esp. HIV and
 PT HPV
 PT HPV
 PS Claim 14; Page 97-98; 172pp; English.
 XX
 CC AAR95965-R95993 represent the nucleic acid recognition units (NAR) of
 CC target binding assemblies (TBA) of the invention. These NARs are
 CC selected from NF-kappa-B, SPI, TATA, human papillomavirus (HPV) E2, HPV
 CC LTR, human immunodeficiency virus (HIV) LTR and Tat binding units. This
 CC sequence represents a SPI sequence NAR. The TBA is recognised by the
 CC target binding region (TBR) of a probe of the invention. The probe of
 CC the invention contains a TBR, a booster binding region (BBR), and an
 CC optional support or attachment (OSA). The TBA contains at least one
 CC NAR, and optionally a linker sequence, an assembly sequence, an
 CC asymmetry sequence, a nuclear localisation signal sequence, and an OSA.
 CC The assembly sequence and asymmetry sequences are responsible for the
 CC folding and association of the NARs. The linker sequence is an
 CC oligopeptide, which does not interfere with the NAR function, but provides
 CC stability and control over the spacing of the NAR from the rest of the
 CC TBA. The OSA is an attached support or indicator, or other means of
 CC localisation of the probe. The probe can be used in a method for
 CC detecting or localising a specific target nucleic acid sequence (TNA).
 CC The method is highly sensitive, and has a high degree of specificity.
 CC The method can be used for detecting specific nucleic acid sequences,
 CC including those found in human cells, in HIV, HPV, and other nucleic acid
 CC containing systems, including bacteria and viruses.
 XX
 SQ Sequence 168 AA;

TD AAU38749 standard; Protein; 173 AA.
 XX AAU38749;
 XX DT 19-DEC-2001 (first entry)
 XX DE Novel subtilisin homologue #6.
 XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KW contact lens cleansing solution; dry cleaning.
 XX Bacillus sp.
 XX OS WO200175087-A2.
 XX PD 11-OCT-2001.
 XX PF 02-APR-2001; 2001WO-US10781.
 XX PR 03-APR-2000; 2000US-194143P.
 XX PA (MAXY-) MAXYGEN INC.
 XX PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Minshull J;
 XX DR WPI; 2001-616689/71.
 XX DR N-PSDB; AAS56604.
 XX PT Novel subtilisin homologue polypeptides having improved endo-protease
 activity relative to mature Savinase subtilisin Polypeptide, useful as
 component of cleaning solutions e.g. laundry detergents, dry cleaning
 XX PS Claim 1; Page 103; 138pp; English.
 XX CC The invention relates to an isolated polypeptide with improved endo-
 protease activity relative to subtilisin homologue polypeptide. Savinase
 (RTM). The invention also relates to an integrated system comprising a
 computer or computer readable medium comprising a database comprising one
 or more sequence records. Each record comprises one or more character
 string corresponding to a nucleic acid or protein sequence of AAS56599-
 AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 comprises a user input interface allowing a user to selectively
 one or more sequence record. The integrated system is useful for
 presenting information pertaining to one of several sequence records
 stored in a database. The method involves determining a list of one or
 more character strings corresponding to the sequence or its subsequence,
 determining which character strings of the list are selected by a user,
 and displaying the selected character strings or aligning the selected
 character string with an additional character string. The method
 further involves displaying an alignment of the selected character
 string with the additional character string, and displaying the list.
 The polynucleotides are useful as probes, primers, sense and antisense
 nucleotides. The subtilisin homologue polypeptides are useful for
 producing antibodies which have diagnostic uses related to the activity,
 distribution and expression of subtilisin homologues. The subtilisin
 homologue polypeptides are used in compositions that serve as cleaning
 solutions in a wide variety of applications including laundry detergents,
 contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 represent the amino acid sequences of novel subtilisin homologues of the
 invention.
 XX Sequence 173 AA;
 XX Query Match 91.7%; Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 SQ 1 gtxxps 6
 1 1 1
 db 71 gtsaps 76
 XX Sequence 173 AA;
 XX Query Match 91.7%; Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query 1 gtxxps 6 Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred No 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 40
 ID AAU38780 standard; Protein; 173 AA.

AC AAU38780;
 XX
 DT 19-DEC-2001 (first entry)
 XX
 Novel subtilisin homologue #37.
 XX
 Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 contact lens cleansing solution; dry cleaning.
 DE
 XX
 OS Bacillus sp.
 OS
 XX
 PN WO200175087-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-US10781.
 XX
 PR 03-APR-2000; 2000US-194143P.
 XX
 PA (MAXY-) MAXYGEN INC.
 XX
 PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Marshull J;
 XX
 DR 2001-616689/71.
 DR N-PSDB; AAS56635.
 XX
 Novel subtilisin homologue polypeptides having improved endo-protease
 activity relative to mature Savinase subtilisin polypeptide, useful as
 component of cleaning solutions e.g. laundry detergents, dry cleaning
 PT
 PT
 PS Claim 1; Page 106; 138pp; English.
 XX
 The invention relates to an isolated polypeptide with improved endo-
 protease activity relative to subtilisin homologue polypeptide. Savinase
 (RTM). The invention also relates to an integrated system comprising a
 computer or computer readable medium comprising a database comprising one
 or more sequence records. Each record comprises one or more character
 string corresponding to a nucleic acid or protein sequence of AAS56599-
 AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 comprises a user input interface allowing a user to selectively
 one or more sequence record. The integrated system is useful for
 presenting information pertaining to one of several sequence records
 stored in a database. The method involves determining a list of one or
 more character strings corresponding to the sequence or its subsequence,
 determining which character strings of the list are selected by a user,
 and displaying the selected character strings or aligning the selected
 character string with an additional character string. The method
 further involves displaying an alignment of the selected character
 string with the additional character string, and displaying the list.
 The polynucleotides are useful as probes, primers, sense and antisense
 nucleotides. The subtilisin homologue polypeptides are useful for
 producing antibodies which have diagnostic uses related to the activity,
 distribution and expression of subtilisin homologues. The subtilisin
 homologue polypeptides are used in compositions that serve as cleaning
 solutions in a wide variety of applications including laundry detergents,
 contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 represent the amino acid sequences of novel subtilisin homologues of the
 invention.
 Sequence 173 AA;
 SQ

Query 1 gtxxps 6 Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred No 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 41
 ID AAU38787
 ID AAU38787 standard; Protein; 173 AA.

XX
 Novel subtilisin homologue #44.
 AC AAU38787;
 XX
 Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 contact lens cleansing solution; dry cleaning.
 DE
 XX
 KW 19-DEC-2001 (first entry)
 XX
 Novel subtilisin homologue #44.
 DE
 XX
 KW Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 contact lens cleansing solution; dry cleaning.
 OS
 XX
 KW
 OS
 XX
 PN WO200175087-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-US10781.
 XX
 PR 03-APR-2000; 2000US-194143P.
 XX
 PA (MAXY-) MAXYGEN INC.
 XX
 PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Marshull J;
 XX
 PR 2001-616689/71.
 DR N-PSDB; AAS56642.
 XX
 Novel subtilisin homologue polypeptides having improved endo-protease
 activity relative to mature Savinase subtilisin polypeptide, useful as
 component of cleaning solutions e.g. laundry detergents, dry cleaning
 PT
 PT
 PS Claim 1; Page 107; 138pp; English.
 XX
 The invention relates to an isolated polypeptide with improved endo-
 protease activity relative to subtilisin homologue polypeptide. Savinase
 (RTM). The invention also relates to an integrated system comprising a
 computer or computer readable medium comprising a database comprising one
 or more sequence records. Each record comprises one or more character
 string corresponding to a nucleic acid or protein sequence of AAS56599-
 AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 comprises a user input interface allowing a user to selectively
 one or more sequence record. The integrated system is useful for
 presenting information pertaining to one of several sequence records
 stored in a database. The method involves determining a list of one or
 more character strings corresponding to the sequence or its subsequence,
 determining which character strings of the list are selected by a user,
 and displaying the selected character strings or aligning the selected
 character string with an additional character string. The method
 further involves displaying an alignment of the selected character
 string with the additional character string, and displaying the list.
 The polynucleotides are useful as probes, primers, sense and antisense
 nucleotides. The subtilisin homologue polypeptides are useful for
 producing antibodies which have diagnostic uses related to the activity,
 distribution and expression of subtilisin homologues. The subtilisin
 homologue polypeptides are used in compositions that serve as cleaning
 solutions in a wide variety of applications including laundry detergents,
 contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 represent the amino acid sequences of novel subtilisin homologues of the
 invention.
 Sequence 173 AA;
 SQ

CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 CC represent the amino acid sequences of novel subtilisin homologues of the
 CC invention.

XX

SQ Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX

Sequence 173 AA;

Qy 1 9ttxxps 6
 ID AAU38804 standard; protein; 173 AA.
 XX
 AAU38804;
 XX
 DT 19-DEC-2001 (first entry)
 XX Novel subtilisin homologue #61.
 DE
 XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KW contact lens cleansing solution; dry cleaning.
 XX
 OS Bacillus sp.
 FN WO200175087-A2.
 XX
 PD 11-OCT-2001.
 XX
 PP 02-APR-2001; 2001WO-US10781.
 XX
 PR 03-APR-2000; 2000US-194143P.
 XX
 PA (MAXYGEN INC.
 XX
 PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Minshull J;
 XX
 WPI; 2001-616689/71.
 DR N-PSDB; AAS56639.
 XX
 Novel subtilisin homologue polypeptides having improved endo-protease
 activity relative to mature Savinase subtilisin polypeptide, useful as
 component of cleaning solutions e.g. laundry detergents, dry cleaning

XX
 PT Claim 1; Page 109; 138pp; English.

XX
 The invention relates to an isolated polypeptide with improved endo-
 protease activity relative to subtilisin homologue polypeptide, Savinase
 (RTM). The invention also relates to an integrated system comprising a
 computer or computer readable medium comprising a database comprising one
 or more sequence records. Each record comprises one or more character
 strings corresponding to a nucleic acid or protein sequence of AAS56599-
 AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 comprises user input interface allowing a user to selectively
 one or more sequence record. The integrated system is useful for
 presenting information pertaining to one of several sequence records
 stored in a database. The method involves determining a list of one or
 more character strings corresponding to the sequence or its subsequence
 determining which character strings of the list are selected by a user,
 and displaying the selected character strings or aligning the selected
 character string with an additional character string. The method
 further involves displaying an alignment of the selected character
 string with the additional character string, and displaying the list.

CC nucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antibodies which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC string with the additional character string, and displaying the list.

CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 CC represent the amino acid sequences of novel subtilisin homologues of the
 CC invention.

XX

SQ Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX

Qy 1 9ttxxps 6
 Db 71 gtssps 76

RESULT 43
 ID AAU38810
 AAU38810 standard; protein; 173 AA.
 XX
 AC AAU38810;
 XX
 DT 19-DEC-2001 (first entry)
 XX Novel subtilisin homologue #67.
 DE
 XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KW contact lens cleansing solution; dry cleaning.
 XX
 OS Bacillus sp.
 XX
 PN WO200175087-A2.
 XX
 PD 11-OCT-2001.
 XX
 PP 02-APR-2001; 2001WO-US10781.
 XX
 PR 03-APR-2000; 2000US-194143P.
 XX
 PA (MAXYGEN INC.
 XX
 PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Minshull J;
 XX
 WPI; 2001-616689/71.
 DR N-PSDB; AAS56635.
 XX
 Novel subtilisin homologue polypeptides having improved endo-protease
 activity relative to mature Savinase subtilisin polypeptide, useful as
 component of cleaning solutions e.g. laundry detergents, dry cleaning

XX
 PT Claim 1; Page 110; 138pp; English.

XX
 The invention relates to an isolated polypeptide with improved endo-
 protease activity relative to subtilisin homologue polypeptide, Savinase
 (RTM). The invention also relates to an integrated system comprising a
 computer or computer readable medium comprising a database comprising one
 or more sequence records. Each record comprises one or more character
 strings corresponding to a nucleic acid or protein sequence of AAS56599-
 AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 comprises user input interface allowing a user to selectively
 one or more sequence record. The integrated system is useful for
 presenting information pertaining to one of several sequence records
 stored in a database. The method involves determining a list of one or
 more character strings corresponding to the sequence or its subsequence
 determining which character strings of the list are selected by a user,
 and displaying the selected character strings or aligning the selected
 character string with an additional character string. The method
 further involves displaying an alignment of the selected character
 string with the additional character string, and displaying the list.

CC nucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antibodies which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC string with the additional character string, and displaying the list.

CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAU38744-AAU38874
 CC represent the amino acid sequences of novel subtilisin homologues of the
 CC invention.

XX

SQ Sequence 173 AA;

Query Match 91.7%; Score 22; DB 22; Length 173;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX

Qy 1 9ttxxps 6
 Db 71 gtssps 76

RESULT 43
 ID AAU38810
 AAU38810 standard; protein; 173 AA.
 XX
 AC AAU38810;
 XX
 DT 19-DEC-2001 (first entry)
 XX Novel subtilisin homologue #67.
 DE
 XX Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KW contact lens cleansing solution; dry cleaning.
 XX
 OS Bacillus sp.
 XX
 PN WO200175087-A2.
 XX
 PD 11-OCT-2001.
 XX
 PP 02-APR-2001; 2001WO-US10781.
 XX
 PR 03-APR-2000; 2000US-194143P.
 XX
 PA (MAXYGEN INC.
 XX
 PI Ness JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;
 PI Minshull J;
 XX
 WPI; 2001-616689/71.
 DR N-PSDB; AAS56635.
 XX
 Novel subtilisin homologue polypeptides having improved endo-protease
 activity relative to mature Savinase subtilisin polypeptide, useful as
 component of cleaning solutions e.g. laundry detergents, dry cleaning

XX
 PT Claim 1; Page 110; 138pp; English.

XX
 The invention relates to an isolated polypeptide with improved endo-
 protease activity relative to subtilisin homologue polypeptide, Savinase
 (RTM). The invention also relates to an integrated system comprising a
 computer or computer readable medium comprising a database comprising one
 or more sequence records. Each record comprises one or more character
 strings corresponding to a nucleic acid or protein sequence of AAS56599-
 AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 comprises user input interface allowing a user to selectively
 one or more sequence record. The integrated system is useful for
 presenting information pertaining to one of several sequence records
 stored in a database. The method involves determining a list of one or
 more character strings corresponding to the sequence or its subsequence
 determining which character strings of the list are selected by a user,
 and displaying the selected character strings or aligning the selected
 character string with an additional character string. The method
 further involves displaying an alignment of the selected character
 string with the additional character string, and displaying the list.

CC nucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antibodies which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC string with the additional character string, and displaying the list.

The polynucleotides are useful as probes, primers, sense and antisense nucleotides. The subtilisin homologue polypeptides are useful for producing antibodies which have diagnostic uses related to the activity, distribution and expression of subtilisin homologues. The subtilisin homologue polypeptides are useful in compositions that serve as cleaning solutions in a wide variety of applications including laundry detergents, contact lens cleansing solutions, and dry cleaning. AU38744-AU38874 represent the amino acid sequences of novel subtilisin homologues of the invention.

Query Match	91.78;	Score 22;	DB 22;	Length 173;
Best Local Similarity	66.78;	Pred. No.	6.3e+02;	
Matches 4; Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;

Y 1 qtxxps 6
| | |
71 qtsaps 76

RESULT 4
AU38835 2 ANU38835 standard; protein: 173 AA.

AAU38835;
19-DEC-2001 (first entry)
Novel subtilisin homologue #92.
Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
contract lens cleaning solution; arr. laundry

X Bacillus sp. Y

WO2000175087-A2. 11-05cm-2001

X 02-APR-2001; 2001WO-US10781.

03-APR-2000; 2000US-194143P.
(MAXY-) MAXYGEN INC.

X Novel subtilisin homologue polypeptides having improved endo-protease activity relative to mature Savinase subtilisin polypeptide, useful as component of cleaning solutions e.g. laundry detergents, dry cleaning N 27, 2001-05-07/1.
N-PSDB; AAS56690.

X 5 Claim 1; Page 113; 138pp; English.

The invention relates to an isolated polypeptide with improved endo-peptidase activity relative to subtilisin homologue polypeptide, Savinase (RTM). The invention also relates to an integrated system comprising a computer or computer readable medium comprising a database comprising one or more sequence records. Each record comprises one or more character string corresponding to a nucleic acid or protein sequence of AAS56599-AAS56728 or AAU38744-AAU38873 respectively. The integrated system comprises a user input interface allowing a user to selectively one or more sequence records. The integrated system is useful for presenting information pertaining to one of several sequence records stored in a database. The method involves determining a list of one or more character strings corresponding to the sequence or its subsequence, determining which character strings of the list are selected by a user,

and displaying the selected character strings or aligning the selected character string with an additional character string. The method further involves displaying an alignment of the selected character string with the additional character string, and displaying the list. The polynucleotides are useful as probes, primers, sense and antisense nucleotides. The subtilisin homologue polypeptides are useful for producing antibodies which have diagnostic uses related to the activity, distribution and expression of subtilisin homologues. The subtilisin homologue polypeptides are used in compositions that serve as cleaning solutions in a wide variety of applications including laundry detergents, contact lens cleansing solutions, and dry cleaning. AAU38744 AAU3874 represent the amino acid sequences of novel subtilisin homologues of the invention.

presenting information pertaining to one of several sequence records stored in a database. The method involves determining a list of one or more character strings corresponding to the sequence or its subsequence, determining which character strings of the list are selected by a user, and displaying the selected character strings or aligning the selected character string with an additional character string. The method further involves displaying an alignment of the selected character string with the additional character string, and displaying the list. The polynucleotides are useful as probes, primers, sense and antisense nucleotides. The *subtilisin* homologue polypeptides are useful for producing antibodies which have diagnostic uses related to the activity, distribution and expression of *subtilisin* homologues. The *subtilisin* homologue polypeptides are used in compositions that serve as cleansing solutions in a wide variety of applications including laundry detergents, contact lens cleansing solutions, and dry cleaning. AU38744 AU38874 represent the amino acid sequences of novel *subtilisin* homologues of the invention.

```

Query Match      91.7%;  Score 22;  DB 22;  Length 173;
Best Local Similarity 66.7%;  Pred. No. 6.3e+02;
Matches 4;  Conservative 0;  Mismatches 0;  Indels 0;
          caps 0;

1 9txxps 6
   | | |
71 9tssps 76


```

AAU38869;
19-DEC-2001 (first entry)
Novel *subtilisin* homologue #126.
Subtilisin; *Savinase*; diagnostic; cleaning solution; laundry detergent;
contact lens cleansing solution; dry cleaning.

Bacillus sp.
WO200175087-A2.
11-OCT-2001.
02-APR-2001; 2001WO-US10781.
03-APR-2000; 2000US-194143P.
(MAXY-) MAXYGEN INC.
NESS, JE, Welch M, Giver LJ, Cherry JR, Borchert TV, Stemmer WPC;

WPI: 2001-616689/71.
N-PSDB: AAS56724.

Novel subtilisin homologue polypeptides having improved endo-protease activity relative to mature Savinase subtilisin polypeptide, useful as component of cleaning solutions e.g. laundry detergents, dry cleaning -

Claim 1; Page 117; 138pp; English.

The invention relates to an isolated polypeptide with improved endo-protease activity relative to subtilisin homologue polypeptide, Savinase (RTM). The invention also relates to an integrated system comprising a computer or computer readable medium comprising a database comprising one or more sequence records. Each record comprises one or more character

string corresponding to a nucleic acid or protein sequence of **AAS56599** CC AAS56728 or **AAU38844** - **AAU38873** respectively. The integrated system CC comprises a user input interface allowing a user to selectively CC one or more sequence record. The integrated system is useful for CC presenting information pertaining to one of several sequence records CC stored in a database. The method involves determining a list of one or CC more character strings corresponding to the sequence or its subsequence CC determining which character strings of the list are selected by a user, CC and displaying the selected character strings or aligning the selected CC character string with an additional character string. The method CC further involves displaying an alignment of the selected character CC string with the additional character string, and displaying the list. CC The polynucleotides are useful as probes, primers, sense and antisense CC nucleotides. The subtilisin homologue polypeptides are useful for CC producing antibodies which have diagnostic uses related to the activity CC distribution and expression of subtilisin homologues. The subtilisin CC homologue polypeptides are used in compositions that serve as cleaning CC solutions in a wide variety of applications including laundry detergents, CC contact lens cleansing solutions, and dry cleaning. **AAU38744** - **AAU38874** CC represent the amino acid sequences of novel subtilisin homologues of CC invention.

Query	Match	91.7%	Score	22	DB	22;	Length	173;
Best	Local Similarity	66.7%	Pred.	No.	6.3e+02;			
Matches	4;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps
Qy	1	9txxps 6						0;
Db	71	gtssps 76						
RESULT 47								
ABB68003	ABB68003	standard;	Protein;	175	AA.			
ID	ID							
XX	XX							
AC	AC							
XX	XX							
DT	26-MAR-2002	(first entry)						
XX	XX							
DE	Drosophila melanogaster	polypeptide	SEQ ID NO	30801.				
XX								
KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.							
KW	Drosophila melanogaster.							
XX	XX							
OS	PN	WO200171042-A2.						
XX	XX							
PD	PD	27-SEP-2001.						
XX	XX							
PF	PF	23-MAR-2001; 2001WO-US09331.						
XX	XX							
PR	PR	23-MAR-2000; 2000US-191617P.						
PR	11-JUL-2000; 2000US-0614150.							
XX	XX							
PA	(PEKE) PE CORP NY.							
XX	XX							
PI	Venter JC, Adams M, Li PWD, Myers EW;							
XX	XX							
DR	WPI; 2001-656850/75.							
DR	N-PSDB; ABL12106.							
XX	XX							
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -							

XX Disclosure: SEQ ID NO 30801; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 175 .AA;

Query Match 91.7%; Score 22; DB 22; Length 175;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 98 gtxtps 103

RESULT 48
 ABB68002 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 30798.
 XX WO2001171042-A2.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS XX
 XX PD 27-SEP-2001.
 XX PR 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX DR WPI; 2001-655860/75.
 XX DR N-PSDB; ABL12105.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure: SEQ ID NO 30798; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 182 AA;

Query Match 91.7%; Score 22; DB 22; Length 182;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 105 gtxtps 110

RESULT 49
 AAM40840 DT 22-OCT-2001 (first entry)

AC AAM40840;
 AC Human polypeptide SEQ ID NO 5771.
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; haemotactic;
 KW angiotoxic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX OS Homo sapiens.
 XX PN WO200153312-A1.
 XX PD 26-JUL-2001.
 XX PF 26-DEC-2000; 2000WO-US34263.
 XX PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0398042.
 PR 19-JUL-2000; 2000US-0320312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue J, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX DR WPI; 2001-442253/47.
 DR N-PSDB; AAI59996.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 2; SEQ ID NO 5771; 10070pp; English.
 PS XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies; Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activ/inhibit activity, chemokinetic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed

CC specification.
 XX Sequence 190 AA;
 SQ

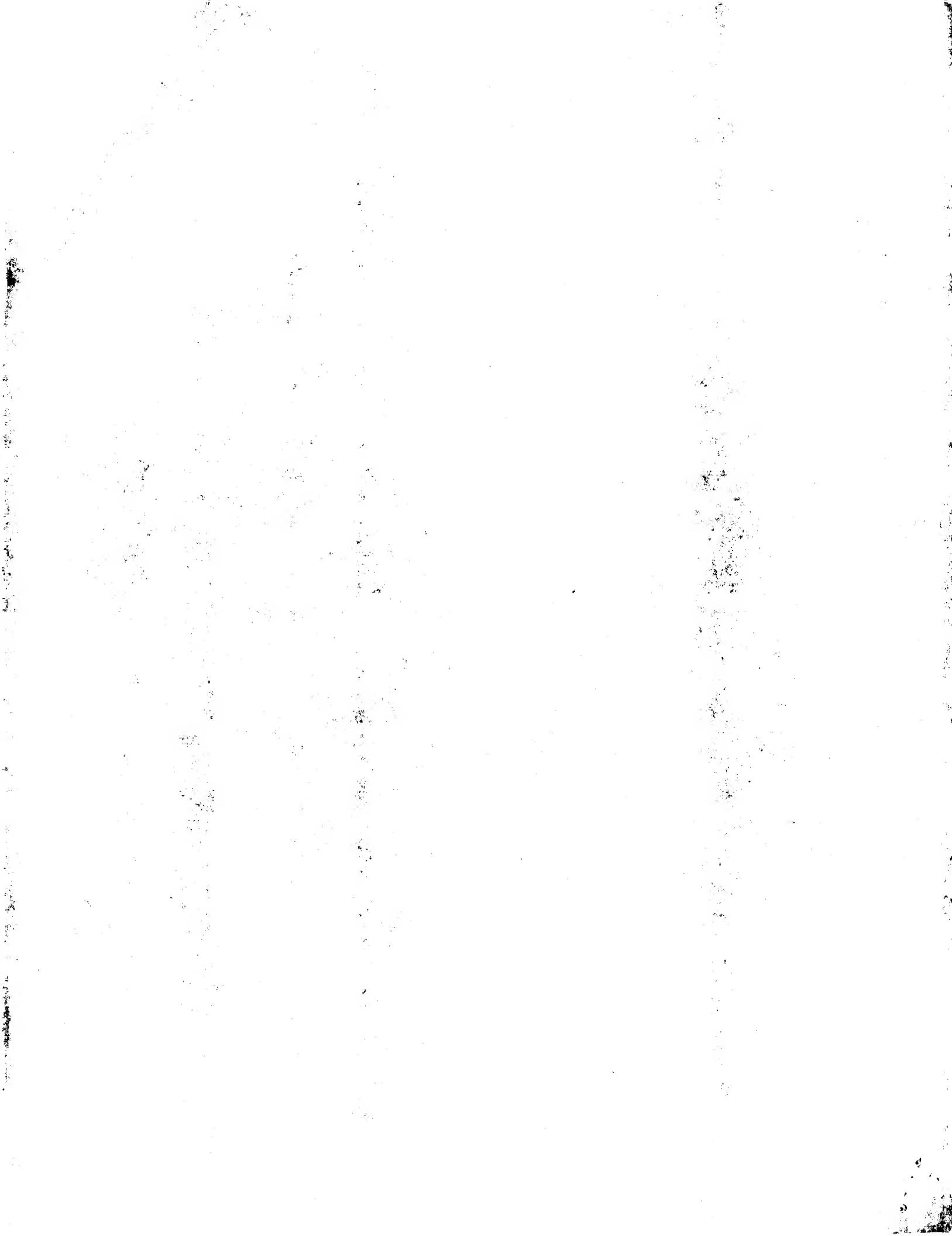
Query Match 91.7%; Score 22; DB 22; Length 190;
 Best Local Similarity 66.7%; Pred. No. 6.8e-02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 9txxps 6
 Db 39 9tstps 44

RESULT 50
 AAU40065 ID AAU40065 standard; Protein: 194 AA.
 XX
 AC AAU40065;
 XX
 DT 13-FEB-2002 (first entry)
 XX Propionibacterium acnes immunogenic protein #961.
 DE XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 FF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORY) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59510.

PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 1260; 1069pp; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC *P. acnes* is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of *P. acnes* in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for *P. acnes* proteins. These antibodies can be used to
 CC downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 Note: The sequence data for this patent did not form part of the printed



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw mode!

Run on: September 3, 2002, 09:25:08 ; Search time 12.98 Seconds
(without alignments)
11.291 Million cell updates/sec

Title: BASK-853-CLAIM5
Perfect score: 24
Sequence: 1 gtxxps 6

Scoring table: BLOSUM62
Gapox 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Issued_Patents_AA:
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:
5: /cgn2_6/ptodata/2/iaa/PCUTS_COH.pep:
6: /cgn2_6/ptodata/2/iaa/backfile1.pep:
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	22	91.7	1	US-08-116-500-41	Sequence 41, Appl	
2	22	91.7	38	1	US-08-471-052A-41	Sequence 41, Appl
3	22	91.7	38	1	US-08-189-331-41	Sequence 41, Appl
4	22	91.7	38	2	US-08-471-939-41	Sequence 41, Appl
5	22	91.7	38	2	US-08-471-800-41	Sequence 41, Appl
6	22	91.7	166	1	US-08-665-617-2	Sequence 41, Appl
7	22	91.7	168	2	US-08-353-476-73	Sequence 41, Appl
8	22	91.7	241	2	US-08-353-476-116	Sequence 41, Appl
9	22	91.7	273	2	US-08-353-476-112	Sequence 41, Appl
10	22	91.7	304	4	US-09-372-422A-40	Sequence 40, Appl
11	22	91.7	1050	4	US-09-428-711A-16	Sequence 40, Appl
12	22	91.7	170	3	US-09-040-786-5	Sequence 40, Appl
13	21	87.5	40	5	PCT-US96-08730-14	Sequence 39, Appl
14	21	87.5	56	2	US-08-537-400-32	Sequence 39, Appl
15	21	87.5	153	4	US-09-228-986-83	Sequence 39, Appl
16	21	87.5	170	2	US-08-483-101-3	Sequence 39, Appl
17	21	87.5	294	3	US-09-081-180-5	Sequence 39, Appl
18	21	87.5	225	1	US-08-290-979A-8	Sequence 38, Appl
19	21	87.5	295	2	US-08-481-956A-9	Sequence 38, Appl
20	21	87.5	295	2	US-08-629-291A-9	Sequence 38, Appl
21	21	87.5	295	2	US-08-650-335B-9	Sequence 38, Appl
22	21	87.5	334	2	US-08-997-080-45	Sequence 45, Appl
23	21	87.5	334	2	US-08-997-362-45	Sequence 45, Appl
24	21	87.5	334	3	US-08-873-970-45	Sequence 45, Appl
25	21	87.5	334	4	US-09-095-855-45	Sequence 45, Appl
26	21	87.5	334	4	US-08-705-347A-41	Sequence 45, Appl
27	21	87.5	334	4	US-08-480-474-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-176-500-41
; Sequence 41, Application US/08176500
; Patent No. 5498538
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/013,416
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-143
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790 9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-176-500-41

Query Match 91.7%; Score 22; DB 1; Length 38;
 Best Local Similarity 66.7%; Pred. No. 68;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 20 GTTSPS 25

RESULT 2
 US-08-471-052A-41
 ; Sequence 41, Application US/08471052A
 ; Patent No. 5625033
 ; GENERAL INFORMATION:
 ; APPLICANT: Kay, B. K.
 ; FOWLKES, D. M.
 ; TITLE OF INVENTION: totally Synthetic Affinity Reagents
 ; NUMBER OF SEQUENCES: 166
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08471052A
 ; FILING DATE: 06-JUNE-1995
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-179
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-3090
 ; TELEFAX: 212 869-8864/9741
 ; INFORMATION FOR SEQ ID NO: 41:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 38 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-189-331-41

Query Match 91.7%; Score 22; DB 1; Length 38;
 Best Local Similarity 66.7%; Pred. No. 68;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 20 GTTSPS 25

RESULT 3
 US-08-189-331-41
 ; Sequence 41, Application US/08189331
 ; Patent No. 5747334
 ; GENERAL INFORMATION:
 ; APPLICANT: Kay, B. K.
 ; FOWLKES, D. M.
 ; TITLE OF INVENTION: totally Synthetic Affinity Reagents
 ; NUMBER OF SEQUENCES: 186
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/189-331
 ; FILING DATE: Concurrently herewith
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-155
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-3090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 41:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 38 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-189-331-41

Query Match 91.7%; Score 22; DB 1; Length 38;
 Best Local Similarity 66.7%; Pred. No. 68;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 20 GTTSPS 25

RESULT 4
 US-08-471-939-41
 ; Sequence 41, Application US/08471939
 ; Patent No. 5844076
 ; GENERAL INFORMATION:
 ; APPLICANT: Kay, B. K.
 ; FOWLKES, D. M.
 ; TITLE OF INVENTION: totally Synthetic Affinity Reagents
 ; NUMBER OF SEQUENCES: 141
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/471,939
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/013,416
 ; FILING DATE: 01-FEB-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-143

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8664/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 38 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 S-08-471-939-41

Query Match 5 Best Local Similarity 91.7%; Score 22; DB 2; Length 38;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 / 1 gtxxps 6
 , 11 11 20 GTTS/PS 25

RESULT 6
 US-08-471-068-41
 ; Sequence 41, Application US/08471068
 ; Patent No. 5948635
 ; GENERAL INFORMATION:
 ; APPLICANT: Fowlkes, D. M.
 ; TITLE OF INVENTION: totally Synthetic Affinity Reagents
 ; NUMBER OF SEQUENCES: 186
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/471,068
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/189,331
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-155
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 41:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 38 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-471-068-41

Query Match 5 Best Local Similarity 91.7%; Score 22; DB 2; Length 38;
 Matches 4; Conservative 0; Mismatches 2; Indels 0;
 / 1 gtxxps 6
 , 11 11 20 GTTS/PS 25

RESULT 7
 US-08-665-617-2
 ; Sequence 2, Application US/08665617
 ; Patent No. 5663316
 ; GENERAL INFORMATION:
 ; APPLICANT: Xudong, Yin
 ; TITLE OF INVENTION: Gene and Protein for Regulation of C
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Salivanchik & Salivanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida

Query Match 5 Best Local Similarity 91.7%; Score 22; DB 2; Length 38;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 / 1 gtxxps 6
 , 11 11 20 GTTS/PS 25

RESULT 8
 US-08-471-800-41
 ; Sequence 41, Application US/08471800
 ; Patent No. 5852167
 ; GENERAL INFORMATION:
 ; APPLICANT: Kay, B. K.
 ; TITLE OF INVENTION: totally Synthetic Affinity Reagents
 ; NUMBER OF SEQUENCES: 141
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/471,800
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/013,416
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-143
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8664/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 41:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 38 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-471-800-41

Page 5

```

Query Match      91.7%;  Score 22;  DB 4;  Length 294;
Best Local Similarity 66.7%;  Pred. No. 4.3e+02;
Matches 4;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

Qy   1  gtxxps 6
      ||| |
Db   158  GTTTPS 163

RESULT 12
US-09-428-711A-16
; Sequence 16, Application US/09428711A
; Patent No. 6358720
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Masaaki
; APPLICANT: Shirasawa, Takuji
; APPLICANT: Tokumitsu, Hiroshi
; APPLICANT: No. 6358720buchi Teruhisa
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE
; FILE REFERENCE: 06501-045001
; CURRENT APPLICATION NUMBER: US/09/428,711A
; CURRENT FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: PCT/JP98/01246
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: JP 9/124798
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1050
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-428-711A-16

Query Match      91.7%;  Score 22;  DB 4;  Length 1050;
Best Local Similarity 66.7%;  Pred. No. 1.4e+03;
Matches 4;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

Qy   1  gtxxps 6
      ||| |
Db   745  GTSSPS 750

RESULT 13
PCT-US96-08/730-14
; Sequence 14, Application PC/TUS9608730
; GENERAL INFORMATION:
; APPLICANT: Cassells, Frederick
; APPLICANT: Anderson, Jeffrey
; APPLICANT: Carter, John Mark
; TITLE OF INVENTION: Methods of Raising Antibodies Against E.
; TITLE OF INVENTION: Coli of the Family CSF-CFA./1
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glenna Hendricks
; STREET: P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: USA
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08730
; FILING DATE: 03-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

```

```

NAME: Hendricks, Glenna
REGISTRATION NUMBER: 3,2,535
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 591-4470
TELEFAX: (703) 591-4428
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
PCT-US96-08730-14

RESULT 14
US-08-537-400-32
; Sequence 32, Application US/08537400
; Patent No. 5939301
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Cloned DNA Polymerases From Thermotoga
; TITLE OF INVENTION: nespolitana And Mutants Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,400
; FILING DATE: 02-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,423
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/370,190
; FILING DATE: 09-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800002
; TELEPHONE: 202-371-2560
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-537-400-32

Query Match 87.5%; Score 21; DB 2; Length 56;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 gtxxps 6
Db 44 GTLSPS 49

RESULT 15
US-09-228-986-83
; Sequence 83, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 83
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pinus radiata
; US-09-228-986-83

Query Match 87.5%; Score 21; DB 4; Length 153;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 gtxxps 6
Db 103 GTLSPS 108

RESULT 16
US-08-483-101-3
; Sequence 3, Application US/08483101
; Patent No. 5932715
; GENERAL INFORMATION:
; APPLICANT: Scott, June R.
; APPLICANT: Froehlich, Barbara
; APPLICANT: Caron, Judy
; TITLE OF INVENTION: CS2 Proteins and Coding Sequences
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,101
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 6-95
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

; TELECOMMUNICATION INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 6-95
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-483-101-3

RESULT 17
Query Match Score 21; DB 2; Length 170;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 9ttxps 6
Db 45 GTALPS 50

RESULT 18
US-09-081-180-5
; Sequence 5, Application US/09081180
; Patent No. 602847
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; ATTORNEY/AGENT INFORMATION:
; ADDRESS: ZymoGenetics
; STREET: 1201 Eastlake Ave. E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040/786
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/041,263
; FILING DATE: March 19, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan E.
; REGISTRATION NUMBER: 41,156
; REFERENCE/DOCKET NUMBER: 97-17
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-040-786-5

Query Match Score 21; DB 3; Length 170;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 9ttxps 6
Db 95 GTYSPS 100

RESULT 19
US-08-290-979A-8
; Sequence 8, Application US/08290979A
; Patent No. 5610046
; GENERAL INFORMATION:
; APPLICANT: VAN OOLJEN, Albert J.H.
; ATTORNEY/AGENT INFORMATION:
; ADDRESS: ZymoGenetics
; STREET: 1201 Eastlake Ave. E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081/180
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/041,263
; FILING DATE: March 19, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan E.
; REGISTRATION NUMBER: 41,156
; REFERENCE/DOCKET NUMBER: 97-17C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-081-180-5

Query Match Score 21; DB 3; Length 170;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 9ttxps 6
Db 95 GTYSPS 100

```

; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/290,979A
 ; FILING DATE: 22-SEP-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KATE H. MURASHIGE
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 4615-0045.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0764
 ; TELEX: 90-4030 MRSN FOERS WSH
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 225 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-290-979A-8

Query Match 87.5%; Score 21; DB 1; Length 225;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 21
 US-08-629-291A-9
 ; Sequence 9, Application US/08629291A
 ; Patent No. 5959174
 ; GENERAL INFORMATION:
 ; APPLICANT: Coruzzi, Gloria
 ; ATTORNEY: Oliveira, Igor
 ; APPLICANT: Lam, Hon-Ming
 ; APPLICANT: Hsieh, Ming Hsuiun
 ; TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/629,291A
 ; FILING DATE: 08-APR-1996
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REPOSITORY/DOCKET NUMBER: 5914-050
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 205 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-629-291A-9

Query Match 87.5%; Score 21; DB 2; Length 295;
 Best Local Similarity 66.7%; Pred. No. 7.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 20
 US-08-481-956A-9
 ; Sequence 9, Application US/08481956A
 ; Patent No. 5824867
 ; GENERAL INFORMATION:
 ; APPLICANT: Coruzzi, Gloria
 ; ATTORNEY: Oliveira, Igor
 ; APPLICANT: Lam, Hon-Ming
 ; APPLICANT: Hsieh, Ming Hsuiun
 ; TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/481,956A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864

```

RESULT 22
US-08-658-335B-9
; Sequence 9, Application US/08658335B
; Patent No. 5981703
; GENERAL INFORMATION:
; APPLICANT: Coruzzi, Gloria
; APPLICANT: Oliveira, Igor
; APPLICANT: Lam, Hon-Ming
; APPLICANT: Hsieh, Ming-Hsien
; TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,335B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 5914-052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9050
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-658-335B-9

Query Match Score 21; DB 2; Length 295;
Best Local Similarity 66.7%; Pred. No. 7.4e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
      |   |
      |   |
      |   |
Db 65 GTVSPSPS 70

RESULT 23
US-08-997-080-45
; Sequence 45, Application US/08997080
; Patent No. 5966524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

```

NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX: 206-269-0563
SEQUENCE CHARACTERISTICS:
LENGTH: 334 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-997-362-45

RESULT 26
US-08-997-362-45
Query Match 87.5%; Score 21; DB 2; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; MOLECULE TYPE: protein
; US-08-873-970-45

Query Match 87.5%; Score 21; DB 3; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; MOLECULE TYPE: protein
; US-08-995-855-45
; Sequence 45, Application US/09095855
; Patent No: 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; NUMBER OF SEQUENCES: Treatment and Diagnosis of Mycobacterial Infections
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-095-855-45

Query Match 87.5%; Score 21; DB 4; Length 334;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
; MOLECULE TYPE: protein
; US-08-995-855-45
; Sequence 45, Application US/09095855
; Patent No: 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; NUMBER OF SEQUENCES: Treatment and Diagnosis of Mycobacterial Infections
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,970
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-095-855-45

```

```

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 45
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
; US-09-324 542 45

; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEES: Specman Picard PLLC
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,347A
; FILING DATE: 28-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; RECOMMUNICATION INFORMATION:
; TELEPHONE: 206.269.0565
; TELEFAX: 206.269.0563
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-705-347A-45

; Query Match 87.5%; Score 21; DB 4; Length 334;
; Best Local Similarity 66.7%; Pred. No. 8.3e+02;
; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; Qy 1 gtxxps 6
; Db 253 GTGTPS 258

; RESULT 29
; US-09-198-955A-10
; Sequence 10, Application US/09198955A
; Patent No. 6187580
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schuelin, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjorrvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schorri, Kirk
; APPLICANT: Kongbak, Lars
; TITLE OF INVENTION: NO. 6187580e1 Pectate Lyases
; CURRENT APPLICATION NUMBER: US/09/198,955A
; FILE REFERENCE: 5378.200-US
; CURRENT FILING DATE: 1998-11-24
; PRIORITY APPLICATION NUMBER: 1343/97
; PRIORITY FILING DATE: 1997-11-24
; PRIORITY APPLICATION NUMBER: 1344/97
; PRIORITY FILING DATE: 1997-11-24
; PRIORITY APPLICATION NUMBER: 60/067,249
; PRIORITY FILING DATE: 1997-12-02
; PRIORITY APPLICATION NUMBER: 60/067,240
; PRIORITY FILING DATE: 1997-12-02
; PRIORITY APPLICATION NUMBER: 09/073,684
; PRIORITY FILING DATE: 1998-05-06
; PRIORITY APPLICATION NUMBER: 09/184,217
; PRIORITY FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 10
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Bacillus sp.
; US-09-198-955A-10

; Query Match 87.5%; Score 21; DB 4; Length 335;
; Best Local Similarity 66.7%; Pred. No. 8.4e+02;
; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; Qy 1 gtxxps 6
; Db 84 GTTTPS 89

; RESULT 30
; US-09-184-217-1
; Sequence 1, Application US/09184217
; Patent No. 6255950
; GENERAL INFORMATION:
; APPLICANT: Lange, Niels E.K.
; APPLICANT: Schuelin, Martin
; APPLICANT: Kongbak, Lars
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schorri, Kirk
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; OF Immunologically Mediated Skin Disorders
; CURRENT APPLICATION NUMBER: US/09/324,542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194

```

APPLICANT: Bjornvad, Mads E.
; APPLICANT: Husain, Philip A.
; TITLE OF INVENTION: Biopreparation Of Textiles At High
; FILE REFERENCE: 5729.000-US
; CURRENT APPLICATION NUMBER: US/09/184,217
; CURRENT FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 335
; TYPE: PRT
; ORGANISM: bacillus sp.
; US-09-184-217-1

Query Match 87.5%; Score 21; DB 4; Length 335;
; Best Local Similarity 66.7%; Pred. No. 8.4e+02;
; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; Gaps 0;
Qy 1 gtxxps 6
Db 84 GTTTPS 89

RESULT 31
US-09-188-930-184
; Sequence 184; Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 110000 1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 184
; LENGTH: 340
; TYPE: PRT
; ORGANISM: mouse
; US-09-188-930-184

Query Match 87.5%; Score 21; DB 4; Length 340;
; Best Local Similarity 66.7%; Pred. No. 8.5e+02;
; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; Gaps 0;
Qy 1 gtxxps 6
Db 120 GTAGPS 125

RESULT 32
US-09-053-866-2
; Sequence 2; Application US/09053866
; Patent No. 6111075
; GENERAL INFORMATION:
; APPLICANT: Xu, Weifeng
; APPLICANT: Presnell, Scott R.
; APPLICANT: Yee, David P.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: PROTEASE-ACTIVATED RECEPTOR
; TITLE OF INVENTION: PAR4 (ZCHEM2)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East

CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leith, Debra K
; REGISTRATION NUMBER: 32,619
; REFERENCE/DOCKET NUMBER: 98-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6674
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-09-053-866-2

Query Match 87.5%; Score 21; DB 3; Length 385;
; Best Local Similarity 66.7%; Pred. No. 9.5e+02;
; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; Gaps 0;
Qy 1 gtxxps 6
Db 19 GTQTPS 24

RESULT 33
US-09-409-199-3
; Sequence 3; Application US/08409199
; Patent No. 5532153
; GENERAL INFORMATION:
; APPLICANT: XU, Shuang-yong
; APPLICANT: XIAO, Jianping
; TITLE OF INVENTION: METHOD FOR CLONING AND
; PRODUCING THE SACI RESTRICTION
; TITLE OF INVENTION: PRODUCING THE SACI RESTRICTION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEW ENGLAND BIOLABS, INC.
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/409,199
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.

```

; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5034; 292
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-409-199-3

RESULT 34
US-09-171-461-32
; Sequence 32, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652_1800000
; CURRENT APPLICATION NUMBER: US/09/171_461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 32
; LENGTH: 439
; TYPE: PRT
; ORGANISM: CELO VIRUS
; FEATURE:
; OTHER INFORMATION: Position: 5366..6685/Product:IVa2
; US-09-171-461-32

Query Match 87.5%; Score 21; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 9 6e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Gaps 0;
Qy 1 gtxxps 6
Db 375 GTKTPS 380

RESULT 35
US-08-956-254-2
; Sequence 2, Application US/08956254A
; Patent No. 6013265
; GENERAL INFORMATION:
; APPLICANT: AURELIAN, LAURE
; TITLE OF INVENTION: Vaccine Composition for Herpes Simplex Virus and
; TITLE OF INVENTION: Methods of Using
; FILE REFERENCE: 14211A
; CURRENT APPLICATION NUMBER: US/08/956_254A
; CURRENT FILING DATE: 1997-10-22
; EARLIER APPLICATION NUMBER: US 60/029, 093
; EARLIER FILING DATE: 1996-10-22
; NUMBER OF SEQ ID NOS: 2

Query Match 87.5%; Score 21; DB 4; Length 439;
Best Local Similarity 66.7%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Gaps 0;
Qy 1 gtxxps 6
Db 19 GTKTPS 24

RESULT 36
US-09-008-388-1
; Sequence 1, Application US/09008388
; Patent No. 6034131
; GENERAL INFORMATION:
; APPLICANT: AURELIAN, LAURE
; TITLE OF INVENTION: VACCINE COMPOSITION FOR HERPES
; TITLE OF INVENTION: SIMPLEX VIRUS AND METHOD OF USING
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PEPPER HAMILTON LLP
; STREET: 600 FOURTEENTH STREET, N.W., Suite 500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005-2004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-WINDOWS
; SOFTWARE: WP 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/008, 388
; FILING DATE: December 2, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ramsey, William S.
; REGISTRATION NUMBER: 32,715
; REFERENCE/DOCKET NUMBER: 114178-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 220-1280
; TELEFAX: (202) 220-1665
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
;
```

FRAGMENT TYPE: INTERNAL
; ORIGIN: INTERNAL SOURCE: INTERNAL
; ORGANISM: HERPES SIMPLEX
; STRAIN: RECOMBINANT
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITES:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; US-09-008-388-1

Query Match 87.5%; Score 21; DB 2; Length 448;
; Best Local Similarity 66.7%; Pred. No. 1.1e+03;
; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
; Db 116 GTSGPS 121

RESULT 38
; US-08-476-509B-2
; Sequence 2, Application US/08476509B
; Patent No. 6,034,212
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; UNITS: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; LOCATION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; THEREOF
; TITLE OF INVENTION:
; TITLE OF INVENTION:
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,509B
; FILING DATE: 01-DEC-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Bsqq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101 CJP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343 1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 448 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-509B-2

RESULT 37
; US-09-015-815-1
; Sequence 1, Application US/09015815
; Patent No. 5,965,356
; GENERAL INFORMATION:
; APPLICANT: AURELTON, LAURE
; APPLICANT: SMITH, CYNTHIA
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS TYPE SPECIFIC SEROASSAY
; CURRENT APPLICATION NUMBER: US/09/015,815
; FILE REFERENCE: 1437LA
; CURRENT FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: US 60/036,622
; EARLIER FILING DATE: 1997-01-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 448
; TYPE: PRT
; ORGANISM: herpes simplex virus-2
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(446)
; OTHER INFORMATION: PROTEIN KINASE DOMAIN OF ICP10 SUBUNIT OF HSV-2
; PUBLICATION INFORMATION:
; AUTHORS: CHUNG ET AL.,
; JOURNAL: J. Virol.,
; VOLUME: 63
; PAGES: 389-3398
; DATE: 1989
; PUBLICATION INFORMATION:
; AUTHORS: NELSON ET AL.,
; JOURNAL: J. Biol. Chem.,
; VOLUME: 271
; PAGES: 17021-17027
; DATE: 1996
; US-09-015-815-1

Query Match 87.5%; Score 21; DB 3; Length 448;
; Best Local Similarity 66.7%; Pred. No. 1.1e+03;
; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
; Db 139 GTLTPS 144

RESULT 39
; US-08-348-518C-2
; Sequence 2, Application US/08348518C
; Patent No. 6,022740
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A

TITLE OF INVENTION: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
 TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
 TITLE OF INVENTION: THEREOF
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klauber & Jackson
 STREET: 411 Hackensack Avenue
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/348,518C
 TELEPHONE: 201-487-5800
 TELEFAX: 201-343-1684
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-101
 TELECOMMUNICATION INFORMATION:
 APPLICATION NUMBER: US/08/348,518C
 TELEPHONE: 201-487-5800
 TELEFAX: 133521
 INQUIRY FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 486 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-348-518C-2

RESULT 40
 US-08-794-795-2
 Sequence 2, Application US/08794795
 ; Patent No. 5916766
 ; GENERAL INFORMATION:
 ; APPLICANT: Elshourbagy, Nabil
 ; APPLICANT: Adamou, John
 ; APPLICANT: Gross, Mitchell
 ; APPLICANT: Lysko, Paul
 ; TITLE OF INVENTION: Human Macro Scavenger Receptor
 ; TITLE OF INVENTION: eptor
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/249,200
 ; FILING DATE: 12-FEB-1999
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/794,795
 ; FILING DATE: 04-FEB-1997
 ; APPLICATION NUMBER: 60/017,699
 ; FILING DATE: 23-MAY-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Prestia, Paul F
 ; REGISTRATION NUMBER: 23,031
 ; REFERENCE/DOCKET NUMBER: ATG-50009-1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-407-0700
 ; TELEX: 846169
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:

LENGTH: 495 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-249-200-2

Query Match 87.5%; Score 21; DB 4; Length 495;
Best Local Similarity 66.7%; Pred. No. 1.2e+03; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 162 GTLAPS 167

RESULT 42
US-09-031-3-92-4
; Sequence 4, Application US/09031392
; Patent No. 5942398

GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26 APR 1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26 FEB 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEX: 617/542-8906
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-299-549-4

RESULT 44
US-09-610-417-4
; Sequence 4, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417

Query Match 87.5%; Score 21; DB 2; Length 534;
Best Local Similarity 66.7%; Pred. No. 1.3e+03; Indels 2; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
Db 87 GTLAPS 92

RESULT 43
US-09-299-549-4
; Sequence 4, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:

```

FILING DATE: 05-Jul-2000
PRIOR APPLICATION DATA:
  FILING DATE: <Unknown>
  APPLICATION NUMBER: 09/299,549
ATTORNEY/AGENT INFORMATION:
  NAME: Meiklejohn, Ph.D., Anita L.
  REGISTRATION NUMBER: 35,283
  RECOMMUNICATION INFORMATION:
    TELEPHONE: 617/542-5070
    TELEX: 617/542-8906
    TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
  LENGTH: 534 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-610-417-4

Query Match          Score 21;  DB 4;  Length 534;
Best Local Similarity 66.7%;  Pred. No. 1.3e+03;  Mismatches 0;  Indels 0;  Gaps 0;
Matches 4;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1  gtxxps 6
      |  |
      |  |
Db      87  GTLAPS 92

RESULT 4 5
US-09-413-814-4
; Sequence 4, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
;   APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
;   APPLICANT: Bristol-Myers Squibb, Co.
;   APPLICANT: Beyer, Stefan
;   APPLICANT: Bloecker, Helmut
;   APPLICANT: Brandt, Petra
;   APPLICANT: Cino, Paul M
;   APPLICANT: Dougherty, Brian A
;   APPLICANT: Goldberg, Steven L
;   APPLICANT: Hofle, Gerhard
;   APPLICANT: Mueller, Joachim
;   APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 1.07
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-4

Query Match          Score 21;  DB 4;  Length 577;
Best Local Similarity 66.7%;  Pred. No. 1.4e+03;  Mismatches 2;  Indels 0;  Gaps 0;
Matches 4;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

Qy      1  gtxxps 6
      |  |
      |  |
Db      371  GTALPS 376

RESULT 4 6

```

US-08-843-530B-28
 ; Sequence 28, Application US/08843530B
 ; Patent No. 59319306

GENERAL INFORMATION:
 ; APPLICANT: Selitrennikoff, Claude
 ; APPLICANT: Agnan, Jacqueline
 ; APPLICANT: Alex, Lisa A.
 ; APPLICANT: Simon, Melvin I.
 ; TITLE OF INVENTION: Osmosensing Histidine Kinases
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Medlen & Carroll, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94104

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #.0, Version #1.30

CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/843,530B
 ; FILING DATE: 16-APR-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Macknight, Kamrin T.
 ; REGISTRATION NUMBER: 38,230
 ; REFERENCE/DOCET NUMBER: UTC-02717
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 28:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 732 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: protein

US-08-843-530B-28

Query Match 87.5%; Score 21; DB 2; Length 47
 ; Best Local Similarity 66.7%; Pred. No. 1.7e+03;
 ; Matches 4; Conservative 0; Mismatches 2; I

Qy	1	gtxxps	6
Db	81	GTLSPS	86

RESULT 47
 US-08-836-943-2
 ; Sequence 2, Application US/08836943
 ; Patent No. 5965391

GENERAL INFORMATION:
 ; APPLICANT: Reinscheid, Dieter
 ; APPLICANT: Eikmanns, Bernhard
 ; APPLICANT: Sahm, Hermann
 ; TITLE OF INVENTION: DNA WHICH REGULATES GENE EXPRESSION
 ; TITLE OF INVENTION: CORYNEFORM BACTERIA
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: The Firm of Karl F. Ross, PC
 ; STREET: 5676 Riverdale Ave.
 ; CITY: Bronx
 ; STATE: New York
 ; ZIP: 10471

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/836,943
 FILING DATE: 08-MAY-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Jonathan
 REGISTRATION NUMBER: 26,963
 REFERENCE/DOCKET NUMBER: 20357
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (718) 884-6600
 TELEFAX: 718/601-1099
 TELEX: 620428
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 739 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-836-943-2

RESULT 48
 Query Match 87.5%; Score 21; DB 2; Length 739;
 Best Local Similarity 66.7%; Pred. No. 1.7e-03;
 Matches 4; Conservative 0; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 716 GTKSPS 721

RESULT 49
 Query Match 87.5%; Score 21; DB 2; Length 774;
 Best Local Similarity 66.7%; Pred. No. 1.8e-03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 292 GTVSFS 297

RESULT 49
 US-08-486-273A-42
 Sequence 42; Application US/08486273A
 ; GENERAL INFORMATION:
 ; PATENT NUMBER: 5985586
 ; PATENT NO. 5985586
 ; GENERAL INFORMATION:
 ; APPLICANT: Daggett, Lorrie P.
 ; APPLICANT: Ellis, Steven B.
 ; APPLICANT: Liaw, Chen W.
 ; APPLICANT: Lu, Chin-Chun
 ; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Brown, Martin, Haller & McClain
 ; STREET: 1660 Union Street
 ; CITY: San Diego
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 92101-2926
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/486,273A
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/231,193
 ; FILING DATE: 20-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie
 ; REGISTRATION NUMBER: 33,779
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-238-0062
 ; TELEFAX: 619-238-0062
 ; INFORMATION FOR SEQ ID NO: 42:
 ; LENGTH: 774 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-486-273A-42

Query Match 87.5%; Score 21; DB 2; Length 774;
 Best Local Similarity 66.7%; Pred. No. 1.8e-03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 292 GTVSFS 297

RESULT 50
US-08-480-474-42
; Sequence 42, Application US/08480474
; Patent No. 6033865.
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Lin, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA ENCODING
; TITLE OF INVENTION: SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, version #1.1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,474
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9382B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-474-42

Query Match 87.5%; Score 21; DB 3; Length 774;
Best Local Similarity 66.7%; Pred. No. 1.8e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 gtxxps 6
Db 292 GIVSPS 297

Search completed: September 3, 2002, 09:25:31
Job time: 23 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2002, 09:25:08 ; Search time 14.73 Seconds
(without alignments)
39.140 Million cell updates/sec

Title: BASK-853-CLAIMS
Perfect score: 24
Sequence: 1 gtxxps 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR_71:
1: pir1:
2: pir2:
3: pir3:
4: pir4:
5: pir5:
6: pir6:
7: pir7:
8: pir8:
9: pir9:
10: pir10:
11: pir11:
12: pir12:
13: pir13:
14: pir14:
15: pir15:
16: pir16:
17: pir17:
18: pir18:
19: pir19:
20: pir20:
21: pir21:
22: pir22:
23: pir23:
24: pir24:
25: pir25:
26: pir26:
27: pir27:
28: pir28:
29: pir29:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	22	91.7	89	2	EB8141		hypothetical prote
2	22	91.7	101	2	S30493		Sp1 protein - mouse
3	22	91.7	105	2	G72515		hypothetical prote
4	22	91.7	111	2	A85695		hypothetical prote
5	22	91.7	135	2	JG2424		hypothetical prote
6	22	91.7	194	2	H90836		hypothetical prote
7	22	91.7	196	2	I39698		blue copper-bindin
8	22	91.7	196	2	T51838		blue copper bindin
9	22	91.7	202	2	B82133		probable outer mem
10	22	91.7	254	2	A81345		probable thioester
11	22	91.7	255	2	EA12725		nitrile hydratase
12	22	91.7	275	2	S47325		myoD protein - zeb
13	22	91.7	280	2	H71320		hypothetical prote
14	22	91.7	285	2	H82356		hypothetical prote
15	22	91.7	289	2	T17957		hypothetical prote
16	22	91.7	316	2	T34553		hypothetical prote
17	22	91.7	319	2	S62196		hypothetical prote
18	22	91.7	379	2	JN0013		synaptic vesicle m
19	22	91.7	381	2	S12223		naringenin-chalcon
20	22	91.7	383	2	E87549		type IV secretion
21	22	91.7	389	1	SYUJCJ		naringenin-chalcon
22	22	91.7	389	2	JC5136		naringenin-chalcon
23	22	91.7	398	2	S412523		naringenin-chalcon
24	22	91.7	410	2	S12224		naringenin-chalcon
25	22	91.7	418	2	F75587		probable glycosylt
26	22	91.7	419	1	SYUJCD		naringenin-chalcon
27	22	91.7	419	2	T36272		hypothetical prote
28	22	91.7	421	2	T41156		probable WD repeat
29	22	91.7	435	2	T24477		hypothetical prote

outer membrane por	30	22	91.7	460	2	S34969
hypothetical prote	31	22	91.7	460	2	T23087
hypothetical prote	32	22	91.7	507	2	T26809
hypothetical prote	33	22	91.7	518	2	FT5460
hypothetical prote	34	22	91.7	522	2	T40520
hypothetical prote	35	22	91.7	542	2	T48488
probable transposa	36	22	91.7	550	2	B70583
hypothetical prote	37	22	91.7	574	2	T29005
unknown protein F2	38	22	91.7	593	2	A96783
hypothetical prote	39	22	91.7	597	2	T35746
dna-k-type molecule	40	22	91.7	629	2	S20516
unknown protein F9	41	22	91.7	651	2	A96781
hypothetical prote	42	22	91.7	691	2	B75622
hypothetical prote	43	22	91.7	693	2	T15152
transcription fact	44	22	91.7	696	2	A26635
protein F27315.24	45	22	91.7	756	2	D96527
regulatory protein	46	22	91.7	788	2	JS7474
hypothetical prote	47	22	91.7	848	2	T20555
hypothetical prote	48	22	91.7	875	2	T20429
capsid protein - 9	49	22	91.7	886	1	A41521
hypothetical prote	50	22	91.7	963	2	T42290

ALIGNMENTS

outer membrane por	RESULT	1				
hypothetical protein PA4033 [imported] - <i>Pseudomonas aeruginosa</i> (strain PA01)	E83141					
C;Species: <i>Pseudomonas aeruginosa</i>						
C;Date: 15-Sep-2000 #sequence_change 31-Dec-2000						
C;Accession: E83141						
R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;						
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.;						
Lory, S.; Olson, M.V.						
Nature 406, 959-964, 2000						
A;Title: Complete genome sequence of <i>Pseudomonas aeruginosa</i> PA01, an opportunistic pa						
A;Reference number: A82950; MUID:20437337						
A;Accession: E83141						
A;Status: preliminary						
A;Molecule type: DNA						
A;Residues: 1-89 <ST0>						
A;Cross-references: GB:AE004820; GB:AE004091; NID:99950223; PIDN:AG07420.1; GSPDB:GN						
A;Experimental source: strain PA01						
C;Genetics:						
A;Gene: PA4033						
Query Match						
Best Local Similarity						
Matches						
4; Conservative						
0; Mismatches						
2; Indels						
0; Gaps						
R;Chester, A.; Charnay, P.						
DNA Seq. 2, 325-327, 1992						
A;Title: Difference in the genomic organizations of the related transcription factors						
C;Date: 13-Jan-1995 #sequence_change 13-Jan-1995						
C;Accession: S30493						
S30493						
sp1 protein - mouse (fragment)						
C;Species: <i>Mus musculus</i> (house mouse)						
C;Date: 13-Jan-1995 #sequence_change 13-Jan-1995						
C;Accession: S30493						
Qy	1	gttxxps	6			
Db	48	GTASPS	53			

Query Match 91.7%; Score 22; DB 2; Length 101;
 Best Local Similarity 66.7%; Pred. No. 81;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 A;Species: Aeropyrum pernix (strain K1)
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999
 C;Accession: G72515
 R;Rawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
 A;Reference number: A72450; MUID:99310339
 A;Accession: G72515
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-105 <RAW>
 A;Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BAA81111.1; PID:d1044897; PID:9510
 A;Experimental source: strain K1
 C;Genetics:
 A;Gene: APE2100

Query Match 91.7%; Score 22; DB 2; Length 105;
 Best Local Similarity 66.7%; Pred. No. 84;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 A;Species: Escherichia coli (strain O157:H7, substrate EDL933)
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001
 C;Accession: A85695
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimambato, E.; Potamousis, K.; Apodaca, N.; Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: A85695
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-111 <STO>
 A;Cross-references: GB:AB005174; NID:912514864; PIDN:AAG56021.1; GSPDB:GN00145; UWGP:219
 A;Experimental source: strain O157:H7, substrate EDL933
 C;Genetics:
 A;Gene: Z1932

Query Match 91.7%; Score 22; DB 2; Length 111;
 Best Local Similarity 66.7%; Pred. No. 89;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 A;Species: Escherichia coli (strain O157:H7, substrate EDL933)
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001
 C;Accession: A85695
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimambato, E.; Potamousis, K.; Apodaca, N.; Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: A85695
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-196 <RES>
 A;Cross-references: EMBL:215058; NID:916202; PIDN:CAA78771.1; PID:916203
 C;Genetics:
 A;Gene: bcb
 A;Introns: 57/3
 C;Superfamily: plastocyanin
 C;Keywords: copper binding; transmembrane protein (His, Cys, His, Glu) #status predicted F;66,107,112,117/Binding site: copper (His, Cys, His, Glu) #status predicted

RESULT 3
 G72515 hypothetical protein APE2100 - Aeropyrum pernix (strain K1)
 C;Species: Aeropyrum pernix
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C;Accession: G72515
 R;Rawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
 A;Reference number: A72450; MUID:99310339
 A;Accession: G72515
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-105 <RAW>
 A;Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BAA81111.1; PID:d1044897; PID:9510
 A;Experimental source: strain K1
 C;Genetics:
 A;Gene: APE2100

Query Match 91.7%; Score 22; DB 2; Length 194;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 A;Species: Escherichia coli (strain O157:H7, substrate EDL933)
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C;Accession: H90836
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C;gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.; DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
 A;Reference number: A99659; MUID:21156231; PMID:11258196
 A;Accession: H90836
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-194 <HAY>
 A;Cross-references: GB:BA000007; PIDN:BAB35087.1; PID:913361128; GSPDB:GN00154
 A;Experimental source: strain O157:H7, substrate RIMD 050952
 C;Genetics:
 A;Gene: Ecs1664

RESULT 6
 H90836 hypothetical protein Ecs1664 [imported] - Escherichia coli (strain O157:H7, substrate EDL933)
 C;Species: Escherichia coli
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C;Accession: H90836
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C;gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.; DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
 A;Reference number: A99659; MUID:21156231; PMID:11258196
 A;Accession: H90836
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-194 <HAY>
 A;Cross-references: GB:BA000007; PIDN:BAB35087.1; PID:913361128; GSPDB:GN00154
 A;Experimental source: strain O157:H7, substrate RIMD 050952
 C;Genetics:
 A;Gene: Ecs1664

Query Match 91.7%; Score 22; DB 2; Length 194;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 A;Species: Escherichia coli (strain O157:H7, substrate EDL933)
 C;Date: 04-Sep-1997 #sequence_revision 04 Sep-1997 #text_change 11-Jun-1999
 C;Accession: C25555
 R;Van Oysel, A.; van Montagu, M.; Inze, D.; Gene 136, 79-85, 1993
 A;Title: A negatively light-regulated gene from Arabidopsis thaliana encodes a protein
 A;Reference number: 139698; MUID:94124044
 A;Accession: 139698
 A;Molecule type: DNA
 A;Residues: 1-196 <RES>
 A;Cross-references: EMBL:215058; NID:916202; PIDN:CAA78771.1; PID:916203
 C;Genetics:
 A;Gene: bcb
 A;Introns: 57/3
 C;Superfamily: plastocyanin
 C;Keywords: copper binding; transmembrane protein (His, Cys, His, Glu) #status predicted F;66,107,112,117/Binding site: copper (His, Cys, His, Glu) #status predicted

RESULT 5
 JC2424 hypothetical 14.3k protein - mouse

F;79-113/Disulfide bonds: #status predicted

Query Match 91.7%; Score 22; DB 2; Length 196;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 158 GTrTPS 163

RESULT 8
 T51838 blue copper binding protein homolog [imported] - *Arabidopsis thaliana*
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Accession: T51838
 C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 17-Nov-2000
 A;Description: Characterization of a wound-inducible *Arabidopsis* gene encoding a protein submitted to the EMBL Data Library, October 1998
 A;Reference number: Z25481
 A;Accession: T51838
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Cross-references: EMBL:Y18227; PIDN:CAA77089.1
 C;Genetics:
 A;Gene: ANI_32
 C;Superfamily: plastocyanin
 C;Keywords: copper binding

Query Match 91.7%; Score 22; DB 2; Length 196;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 158 GTrTPS 163

RESULT 9
 B82133 probable outer membrane lipoprotein Slp VC1987 [imported] - *Vibrio cholerae* (strain N169
 C;Species: *Vibrio cholerae*
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 23-Mar-2001
 C;Accession: B82133
 R;Heideberg, J.F.; Eisen, R.A.; Clayton, W.C.; Nelson, W.C.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermoljeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
 L., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 A;Title: 406, 477-483, 2000
 A;Cross-references: GB:AE004273; GB:AE003852; NID:99656517; PIDN:AAF95135.1; GSPDB:GN001
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VC1987
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-202 <HBI>
 A;Accession: B82133

Query Match 91.7%; Score 22; DB 2; Length 202;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 134 GTTAPS 139

RESULT 10
 A83345 probable thioesterase PA2411 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C;Species: *Pseudomonas aeruginosa*
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C;Accession: A83345
 R;Stover, C.K.; Phan, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.
 Lory, S.; Olson, M.V.
 Nature 406, 955-964, 2000
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
 A;Reference number: A82950; MUID:20437337
 A;Accession: A83345
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-254 <STO>
 A;Cross-references: GB:AE004668; GB:AE004091; NID:9948446; PIDN:AAQ05799.1; GSPDB:GN
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: PA2411
 C;Superfamily: oleoyl-[acyl-carrier-protein] hydrolase; oleoyl-lacyl-carrier-protein

Query Match 91.7%; Score 22; DB 2;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 108 GTAAPS 113

RESULT 11
 E42725 nitrile hydratase 3'-hypothetical protein orfE - *Pseudomonas chlororaphis* (strain B23
 C;Species: *Pseudomonas chlororaphis*
 C;Accession: E42725
 R;Nishiyama, M.; Horinouchi, S.; Kobayashi, M.; Nagasawa, T.; Yamada, H.; Beppu, T.
 J. Bacteriol. 173, 2465-2472, 1991
 A;Title: Cloning and characterization of genes responsible for metabolism of nitrile
 A;Reference number: A42725
 A;Accession: E42725
 C;Accession: E42725
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-255 <NLS>
 A;Cross-references: GB:D90216; NID:9216850; PIDN:BA14248.1; PID:di1014953; PID:q21602

Query Match 91.7%; Score 22; DB 2; Length 255;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 103 GTAAPS 108

RESULT 12
 S47325 myoD protein - zebra fish
 C;Species: *Brachydanio rerio* (zebra fish)
 C;Accession: S47325
 R;Weinberg, E.S.
 submitted to the EMBL Data Library, September 1994
 A;Description: Developmental regulation of zebrafish myo D in wild type, no tail, and
 A;Reference number: S47325
 A;Accession: S47325
 A;Status: preliminary
 A;Molecule type: mRNA

A;Residues: 1-275 <WEI>
 A;Cross-references: EMBL:Z36945; NID:9535814; PID:9535815
 C;Superfamily: human myogenin

RESULT 15
 T17957
 hypothetical protein A454L - Chlorella virus PBCV-1
 C;Species: Chlorella virus PBCV-1
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C;Accession: T17957
 R;Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL Data Library, May 1999
 A;Reference number: Z18806
 A;Accession: T11957
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-289 <GRA>
 A;Cross-references: EMBL:U42580; PIDN: AAC68222_1
 A;Experimental source: specific host Chlorella strain NC64
 C;Genetics:
 A;Note: A454L
 C;Superfamily: Chlorella virus PBCV-1 hypothetical protein A454L

Query Match 91.7%; Score 22; DB 2; Length 275;
 Best Local Similarity 66.7%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 253 GTTAPS 258

RESULT 13
 H71320
 hypothetical protein TP0462 - syphilis spirochete
 C;Species: treponema pallidum subsp. pallidum (syphilis spirochete)
 C;Accession: H71320
 C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
 R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDC
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A;Reference number: A71250; MUID:9832770
 A;Accession: H71320
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-280 <C0J>
 A;Cross-references: GB:AE000520; GB:AE001223; GB:AE000520; NID:93322745; PIDN: AAC65455.1; PID:9332275
 A;Experimental source: strain Nichols
 C;Genetics:
 A;Gene: TP0462

RESULT 16
 T34553
 hypothetical protein DKF2P434L1435.1 - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 03-Nov-2000
 C;Accession: T34553
 R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, October 1999
 A;Reference number: Z21140
 A;Accession: T34553
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-316 <POU>
 A;Cross-references: EMBL:AL122037
 A;Experimental source: adult testis; clone DKF2P434L1435
 C;Genetics:
 A;Note: DKF2P434L1435.1
 C;Superfamily: valine-tRNA ligase

Query Match 91.7%; Score 22; DB 2; Length 280;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 116 GTTAPS 121

RESULT 14
 H85356
 hypothetical protein At4g30510 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Accession: H85356
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A;Reference number: A85001; MUID:20083488
 A;Accession: H85356
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-285 <STO>
 A;Cross-references: GB:NC_001268; NID:97269952; PIDN: CAB79769.1; GSPDB:GN00140
 A;Gene: At4g30510
 A;Map Position: 4

RESULT 17
 S62196
 hypothetical protein 2 - Methanosarcina barkeri
 C;Species: Methanosarcina barkeri
 C;Accession: S62196
 C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 22-Oct-1999
 R;Vorholt, J.A.; Vaupel, M.; Thauer, R.K.
 Eur. J. Biochem. 236, 309-317, 1996
 A;Title: A polyferredoxin with eight [4Fe-4S] clusters as a subunit of molybdenum for
 A;Reference number: S62194; MUID: 96184912
 A;Accession: S62196
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-319 <VOR>

Query Match 91.7%; Score 22; DB 2; Length 285;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 233 GTSSPS 238

A;Cross-references: EMBL:X93084; NID:91124956; PIDN:CAA63627.1; PID:e212293; PID:9112495
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

Query Match 91.7%; Score 22; DB 2; Length 319;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 66 GTSSPS 71

RESULT 18

JN0013 synaptic vesicle membrane protein VAT-1 - Pacific electric ray
 C;Species: Torpedo californica (Pacific electric ray)
 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 31-Mar-2000
 A;Accession: JN0013
 R;Linial, M.; Miller, K.; Scheller, R. H.
 Neuron 2, 1265-1273, 1989

A;Title: VAT-1: an abundant membrane protein from torpedo cholinergic synaptic vesicles.
 A;Reference number: JN0013; MUID:90166593
 A;Molecule type: mRNA
 A;Residues: 1-379 <LIN>
 C;Comment: Synaptic vesicles are responsible for regulating the storage and release of n
 C;Keywords: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
 F;52-350;Domain: Long-chain alcohol dehydrogenase homology <LADH>
 F;67,127,147;Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;273/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 91.7%; Score 22; DB 2; Length 379;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 194 GTASPS 199

RESULT 19

S12223 naringenin-chalcone synthase (EC 2.3.1.74) 1 - tomato
 C;Species: Lycopersicon esculentum (tomato)
 C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-May-2000
 A;Accession: S12223
 R;O'Neill, S. D.; Tong, Y.; Spoerlein, B.; Forkmann, G.; Yoder, J. I.
 Mol. Gen. Genet. 224, 279-288, 1990
 A;Reference number: S12223; MUID:91117196
 A;Map position: 9
 A;Molecule type: mRNA
 A;Residues: 1-381 <ONE>
 C;Superfamily: chalcone synthase
 C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 91.7%; Score 22; DB 2; Length 381;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 23 GTSTPS 28

RESULT 20

E87549 type IV secretion system protein B10, probable [imported] - Caulobacter crescentus
 C;Species: Caulobacter crescentus
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C;Accession: E87549
 R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laiub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Ventler, J.C.; Fraser, C
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of Caulobacter crescentus.
 A;Reference number: A87249; MUID:11259647
 A;Accession: E87549
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-383 <STO>
 A;Cross references: GB:AE005673; PIDN:AAK24393.1; GSPDB:GN00148
 C;Genetics:
 A;Gene: CC2422
 C;Superfamily: tumor-inducing plasmid ptIC58 virB10 protein

Query Match 91.7%; Score 22; DB 2; Length 383;
 Best Local Similarity 66.7%; Pred. No. 2.0e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 74 GTASPS 79

RESULT 21

SYPJCJ naringenin-chalcone synthase (EC 2.3.1.74) J - garden petunia
 N;Alternate names: chalcone synthase
 C;Species: Petunia x hybrida (garden Petunia)
 C;Date: 30-Sep-1991 #sequence_revision 17-Mar-2000 #text_change 05-May-2000
 C;Accession: D72821; JS0309
 R;Koes, R.E.; Specht, C.E.; van den Elzen, P.J.M.; Mol, J.N.M.
 Gene 81, 245-257, 1989
 A;Title: Cloning and molecular characterization of the chalcone synthase multigene fa
 A;Reference number: JS0309; MUID:9034197
 A;Accession: D72821
 A;Molecule type: DNA
 A;Residues: 1-389 <KOE>
 A;Cross-references: EMBL:X14597; NID:920535; PIDN:CAA32737.1; PID:920536
 A;Experimental source: strain violet 30, leaf
 A;Accession: JS0309
 A;Molecule type: DNA
 A;Residues: 1-50, 'D' 52-74, 'V' 76-228, 'I' 230-297, 'L' 299-389 <KOE>
 A;Note: the sequence is revised in GenBank entry PHCSJ release 114, (PIDN:CAA32737.1
 C;Comment: this enzyme plays a central role in the biosynthesis of all classes of fla
 C;Genetics:
 A;Gene: chsJ
 A;Map position: V
 A;Note: chsJ is expressed in various floral tissues and UV illuminated seedlings
 C;Superfamily: chalcone synthase
 C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 91.7%; Score 22; DB 1; Length 389;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 23 GTATPS 28

RESULT 22

JC136 naringenin-chalcone synthase (EC 2.3.1.74) 2 - potato
 N;Alternate names: chalcone synthase; CHS
 C;Species: Solanum tuberosum (potato)

C;Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 05-May-2000
 C;Accession: JC5136; PC4229
 R;Jeon, J.-H.; Kim, H.-S.; Choi, K.H.; Joung, Y.H.; Joung, H.; Byun, S.M.
 Biosci. Biotechnol. Biochem. 60, 1907-1910, 1996
 A;Title: Cloning and characterization of one member of the chalcone synthase gene family
 A;Reference number: JC5136; MUID:97141614
 A;Accession: PC4229
 A;Cross-references: GB:U47738; NID:91470059; PIDN:AAB05239.1; PID:g1470060
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-389 <RE01>
 A;Cross-references: GB:U47738; NID:91470059; PIDN:AAB05239.1; PID:g1470060
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 158-165-367-373 <JE02>
 C;Comment: This enzyme is important in the biosynthesis of all classes of flavonoids in
 C;Superfamily: chalcone synthase; coenzyme A
 C;Keywords: acyltransferase; coenzyme A

RESULT 25
 F75587
 probable glycosyltransferase - Deinococcus radiodurans (strain R1)
 C;Species: Deinococcus radiodurans
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C;Accession: F75587
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vanathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.M.
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science, 286, 1571-1577, 1999
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A;Reference number: A75250; MUID:20036696
 A;Accession: F75587
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-418 <WH1>
 A;Cross-references: GB:AE001825; NID:g6460670; PIDN:AAF12451.1; PID:g646
 A;Experimental source: strain R1
 C;Genetics:
 A;Gene: DRA0329
 A;Map position: 2
 C;Superfamily: glycosyltransferase

Query Match 91.7%; Score 22; DB 2; Length 389;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 23 GTATPS 28

RESULT 26
 SYP0CD
 naringenin-chalcone synthase (EC 2.3.1.74) - parsley
 N;Alternate names: chalcone synthase
 N;Species: Petroselinum crispum (parsley)
 C;Accession: S42523
 C;Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 05-May-2000
 C;Accession: S42523
 A;Title: Coding and 3' non-coding nucleotide sequence of chalcone synthase mRNA and assi
 A;Reference number: S42523
 A;Accession: S42523
 A;Molecule type: mRNA
 A;Residues: 1-398 <RE1>
 A;Cross-references: EMBL:Y01538; NID:g20513; PIDN:CAA24779.1; PID:g20514
 A;Note: the source is designated as Petroselinum hortense
 C;Superfamily: chalcone synthase; coenzyme A
 C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 91.7%; Score 22; DB 2; Length 398;
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 28 GTATPS 33

RESULT 26
 SYP0CD
 naringenin-chalcone synthase (EC 2.3.1.74) D - garden petunia
 N;Alternate names: chalcone synthase
 C;Species: Petunia x hybrida (garden Petunia)
 C;Accession: 30-Sep-1991 #sequence_revision 17-Mar-2000 #text_change 21-Jul-2000
 C;Accession: A7821; JS0312
 R;Kos, R.E.; Spell, C.E.; van den Elzen, P.J.M.; Mol, J.N.M.
 Gene 81, 245-257, 1989
 A;Title: Cloning and molecular characterization of the chalcone synthase multigene fa
 A;Reference number: JS0308; MUID:90034197
 A;Accession: A7821
 A;Molecule type: DNA
 A;Residues: 1-419 <KOE>
 A;Cross-references: GB:X14593; NID:g20528; PIDN:CAA32733.1; PID:g20529
 A;Experimental source: strain Violet 30, leaf
 A;Accession: JS0312
 A;Molecule type: DNA
 A;Residues: 1-117, 'D', 119, 'H', 121-313, 'DI', 316-419 <KO2>
 A;Note: the sequence was revised in GenBank entry PR05HSD, release 114, (PIDN:CAA32733
 C;Comment: This enzyme plays a central role in the biosynthesis of all classes of fla
 C;Genetics:
 A;Gene: chsd
 A;Map position: V
 C;Superfamily: chalcone synthase
 C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

RESULT 24
 S12224
 naringenin-chalcone synthase (EC 2.3.1.74) 2 - tomato
 C;Species: Lycopersicon esculentum (tomato)
 C;Accession: S12224
 C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-May-2000
 R;O'Neill, S.D.; Tong, Y.; Spoerlein, B.; Forkmann, G.; Yoder, J.I.
 Mol. Gen. Genet. 224, 279-288, 1990
 A;Title: Molecular genetic analysis of chalcone synthase in *Lycopersicon esculentum* and
 A;Reference number: S12223; MUID:9117.96
 A;Accession: S12224
 A;Molecule type: mRNA
 A;Residues: 1-410 <ONE>
 C;Genetics:

Query Match 91.7%; Score 22; DB 1; Length 419;
 Best Local Similarity 66.7%; Pred. No. 3e+02; 2; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; A; Gaps 0;
 C;Species: Streptomyces coelicolor

Qy 1 gtxxps 6
 Db 23 GTATPS 28

RESULT 27

T36272 hypothetical protein SCE68.22 - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T36272
 R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A;Reference number: Z21576
 A;Accession: T36272
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: DNA
 A;Residues: 1-419 <MUR>
 A;Cross-references: EMBL:AL079345; PIDN:CAA45358.1; GSPDB:GN00070; SCOEDB:SCE68.22
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOEDB:SCE68.22

Query Match 91.7%; Score 22; DB 2; Length 419;
 Best Local Similarity 66.7%; Pred. No. 3e+02; 2; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; A; Gaps 0;
 C;Species: Schizosaccharomyces pombe

Qy 1 gtxxps 6
 Db 380 GTASPS 385

RESULT 30

T41156 probable WD repeat transcription regulation protein - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999
 C;Accession: T41156
 R;Hilbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A;Reference number: Z21973
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: DNA
 A;Residues: 1-421 <HIL>
 A;Cross-references: EMBL:AL031907; PIDN:CAA21427.1; GSPDB:GN00068; SPDB:SPCC18.13
 A;Experimental source: strain 972h-; cosmid c18
 C;Genetics:
 A;Gene: SPDB:SPCC18.13
 A;Map position: 3
 A;Introns: 227/2

Query Match 91.7%; Score 22; DB 2; Length 421;
 Best Local Similarity 66.7%; Pred. No. 3e+02; 2; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; A; Gaps 0;
 C;Species: Caenorhabditis elegans

Qy 1 gtxxps 6
 Db 410 GTSSPS 415

RESULT 31

T24477 hypothetical protein H13N06.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
 C;Accession: T24477
 R;Maris, B.
 submitted to the EMBL Data Library, August 1996
 A;Reference number: Z19896
 A;Accession: T24477
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: DNA
 A;Residues: 1-435 <WIL>
 A;Cross-references: EMBL:Z88200; PIDN:CAA01584.1; GSPDB:GN00023; CESP:T04H1.7
 A;Experimental source: clone T04H1
 C;Genetics:
 A;Gene: CESP:T04H1.7
 A;Map position: 5
 A;Introns: 54/3; 82/3; 193/1; 239/3; 282/3; 360/1
 C;Superfamily: glucuronosyltransferase

Query Match 91.7%; Score 22; DB 2; Length 435;
 Best Local Similarity 66.7%; Pred. No. 3.1e+02; 2; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; A; Gaps 0;
 C;Species: Pseudomonas aeruginosa (strain S34969)

Qy 1 gtxxps 6
 Db 300 GTTTPS 305

RESULT 30

S34969 outer membrane porin OprE precursor PA0291 [imported] - Pseudomonas aeruginosa (strain S34969)

C;Species: Pseudomonas aeruginosa

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 15-Jun-2001

C;Accession: S34969; E83608

R;Yamano, Y.; Nishikawa, T.; Komatsu, Y.
 Mol. Microbiol. 8, 993-1004, 1993

A;Title: Cloning and nucleotide sequence of anaerobically induced porin protein El (OprE) of Pseudomonas aeruginosa PA01, an opportunistic pathogen

A;Reference number: S34969; MUID:93360827

A;Accession: S34969
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-460 <YAM>

A;Cross-references: GB: D12711; PIDN:BA02207.1; PID:9433418

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; adian, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L'vov, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen

A;Reference number: A82950; MUID:20437337

A;Accession: E83608

A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-460 <STO>

A;Cross-references: GB: AEP004467; GB: AE004091; PIDN:9946133; GSPDB:GN

A;Experimental source: strain PA01

C;Genetics:
 A;Gene: OprE; PA0291
 C;Superfamily: Pseudomonas aeruginosa outer membrane porin PA0958

Query Match 91.7%; Score 22; DB 2; Length 460;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02; 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 54 GTASPS 59

RESULT 31

T23087 hypothetical protein H13N06.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T23087
 R:Lennard, N.
 submitted to the EMBL Data Library, October 1997

A:Accession: T23087
 A:Status: preliminary; translated from GB/EMBL/DBDJ

A: Molecule type: DNA
 A:Residues: 1-460 <WIL>
 A:Cross-references: EMBL:299942; PIDN:CA17068..1; GSPDB:GN00028; CESP:HI3N06.2
 A:Experimental source: clone HI3N06
 C:Genetics:
 A:Gene: CESP:HI3N06.2
 A:Map position: X
 A:Introns: 17/1; 294/1; 331/3; 364/1; 414/2; 432/2

Query Match 91.7%; Score 22; DB 2; Length 460;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 86 GTSTPS 91

RESULT 32
 T26809
 hypothetical protein Y66A7A.8 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999
 R:Steward, C.
 submitted to the EMBL Data Library, October 1998
 A:Accession: T26809
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A: Molecule type: DNA
 A:Residues: 1-507 <WIL>
 A:Cross-references: EMBL:AL032627; PIDN:CAA21552..1; CESP:Y66A7A.8
 R:Matthews, L.
 submitted to the EMBL Data Library, October 1998
 A:Accession: T27293
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A: Molecule type: DNA
 A:Residues: 1-507 <W12>
 A:Cross-references: EMBL:AL032622; PIDN:CAA21503..1; GSPDB:GN00021; CESP:Y66A7A.8
 A:Experimental source: clone Y66A7A
 C:Genetics:
 A:Gene: CESP:Y66A7A.8
 A:Map Position: 3
 A:Introns: 29/3; 93/1; 145/2; 314/3; 358/3

Query Match 91.7%; Score 22; DB 2; Length 507;
 Best Local Similarity 66.7%; Pred. No. 3.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 425 GTSSPS 430

RESULT 33
 F75460
 hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 A:Accession: F75460
 R:White, O.; Eisen, J.A.; Heldelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shem, M.; Vaithhevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma-
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: 219673
 A:Accession: F75460
 A:Status: preliminary
 A: Molecule type: DNA
 A:Residues: 1-518 <WHL>
 A:Cross-references: GB:AE001943; GB:AE000513; NID:96458624; PIDN:AAF10486..1; PID:9645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR903
 A:Map position: 1

Query Match 91.7%; Score 22; DB 2; Length 518;
 Best Local Similarity 66.7%; Pred. No. 3.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 179 GTTTS 184

RESULT 34
 T40520
 hypothetical protein SPBC530.04 - fission yeast (*Schizosaccharomyces pombe*)
 C:Species: *Schizosaccharomyces pombe*
 C:Date: 03-Dec-1999 #sequence_change 03-Dec-1999
 C:Accession: T40520
 R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: 221934
 A:Accession: T40520
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A: Molecule type: DNA
 A:Residues: 1-522 <LYN>
 A:Cross-references: EMBL:AL023634; PIDN:CAA19170..1; GSPDB:GN00067; SPDB:SPBC530.04
 A:Experimental source: strain 972h-; cosmid C530
 C:Genetics:
 A:Gene: SPDB:SPBC530.04
 A:Map position: 2
 A:Introns: 20/3

Query Match 91.7%; Score 22; DB 2; Length 522;
 Best Local Similarity 66.7%; Pred. No. 3.7e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 41 GTSAPS 46

RESULT 35
 T48488
 hypothetical protein T28J14..110 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse ear cress)
 C:Date: 20-Apr-2000 #sequence_change 20-Apr-2000
 C:Accession: T48488
 R:Bayan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.;
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: 224493
 A:Accession: T48488
 A:Status: preliminary
 A: Molecule type: DNA
 A:Residues: 1-542 <BEV>
 A:Cross-references: EMBL:AL163652
 A:Experimental source: cultivar Columbia; BAC clone T28J14
 C:Genetics:
 A:Gene: 5
 A:Map position: 162/3
 A:Introns: 227/1; 251/1; 355/3; 377/2; 397/3; 419/3; 449/2; 487/
 A:Note: T28J14..110

Query Match 91.7%; Score 22; DB 2; Length 542;
 Best Local Similarity 66.7%; Pred. No. 3.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 400 GTATPS 405

RESULT 36

B70583 probable transposase for IS1535 - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whithead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987
 A;Accession: B70583
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Residues: 1-550 <COL>
 A;Cross-references: GB:Z05210; GB:AI123456; NID:91261757; PIDN:CAB08504.1; PIDN:e315222;
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: Ry0922

Query Match 91.7%; Score 22; DB 2; Length 550;
 Best Local Similarity 66.7%; Pred. No. 3.9e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 497 GTTAPS 502

RESULT 37

T29005 hypothetical protein ZK328.3 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T29005
 R;Favello, A.
 A;Description: The sequence of *C. elegans* cosmid ZK328.
 A;Reference number: Z20552
 A;Accession: T29005
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Residues: 1-574 <PAV>
 A;Cross-references: EMBL:U50133; PIDN:AAA91246.1; CESP:ZK328.3
 A;Experimental source: strain Bristol N2
 C;Genetics:
 A;Gene: CESP:ZK328.3
 A;Introns: 63/1; 140/2; 419/1; 436/2; 488/2; 516/1; 545/2

Query Match 91.7%; Score 22; DB 2; Length 574;
 Best Local Similarity 66.7%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 284 GTSTPS 289

RESULT 40

S20516 dna-k⁺ molecular chaperone hsp70, chloroplast - chromophytic alga (*Pavlova lutheri*)
 N;Alternate names: heat shock protein 70
 C;Species: chloroplast *Pavlova lutheri*
 C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
 C;Accession: S20516
 R;Scaramuzzi, C.D.; Stokes, H.W.; Hiller, R.G.
 Plant Mol. Biol. 18, 467-476, 1992
 A;Title: Heat shock Hsp70 protein is chloroplast-encoded in the chromophytic alga *Pav*

A;Reference number: S20516; MUID:92163012
 A;Accession: S20516
 A;Molecule type: DNA
 A;Residues: 1-629 <S>A>
 A;Cross-references: EMBL:X59555; NID:920904; PIDN:CAA42154.1; PID:920905
 C;Genetics:
 A;Gene: hsp70
 A;Genome: chloroplast
 C;Function:
 A;Description: involved in protein folding and assembling/disassembling of protein complex
 C;Superfamily: heat shock protein 70
 C;Keywords: ATP; chloroplast; molecular chaperone

Query Match 91.7%; Score 22; DB 2; Length 629;
 Best Local Similarity 66.7%; Pred. No. 4.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 33 GTTTPS 38

RESULT 41

A96781 hypothetical protein F9E10.5 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Accession: A96781
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;References: A;Bcker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, R.; Theologis, A.; Chung, M.K.; Conn, L.; Conway, A.R.; Creasy, T.H.; Devar, K.; Chin, C.W.; Hughes, B.; Huijar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Kheykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maiti, R.; Marziali, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016/19
 A;Accession: A96781
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-651 <STO>
 A;Cross-references: GB:AE005173; NID:96646755; PIDN:AAF21067.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: F9E10.5
 A;Map Position: 1

Query Match 91.7%; Score 22; DB 2; Length 651;
 Best Local Similarity 66.7%; Pred. No. 4.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 129 GTTSPS 134

RESULT 42

B75622 hypothetical protein - Deinococcus radiodurans (strain R1)
 C;Species: Deinococcus radiodurans
 C;Accession: B75622
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C;References: J.A.; Heidelberger, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; R.; White, O.; Eisen, J.A.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma, M.; Shen, M.; Vamathevan, J.J.; Senger, G.; Mockridge, I.; Sansau, P.; Ruddy, S.; Dudley, K.; Sheer, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A;Title: Genome sequence of the radiotolerant bacterium Deinococcus radiodurans R1.
 A;Reference number: A75250; MUID:20036896
 A;Accession: B75622
 A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 569-598 <RAG>
 A;Experimental source: T-cell line CEM
 A;Note: sequence extracted from NCBI backbone (NCBIP:125980)
 C;Genetics:
 A;Gene: GDB:SP1
 A;Cross-references: GDB:127453; OMIM:18906
 A;RAP position: 19q13.1-19q13.3
 C;Keywords: DNA binding; transcription regulation; zinc finger

Query Match 91.7%; Score 22; DB 2; Length 798;
 Best Local Similarity 66.7%; Pred. No. 5.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 Db 739 GATAPS 744

RESULT 47

T28055
 hypothetical protein ZK858.1 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C;Accession: T28055
 R;White, S.
 Submitted to the EMBL Data Library, September 1996
 A;Cross references: EMBL:Z9759; PIDN:CAB02138.1; GSPDB:GN00019; CESP:ZK858.1
 A;Accession: T28055
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-848 <WIL>
 A;Cross references: EMBL:Z9759; PIDN:CAB02138.1; GSPDB:GN00019; CESP:ZK858.1
 A;Experimental source: clone ZK858
 C;Genetics:
 A;Gene: CESP:ZK858.1
 A;Map position: 1
 A;Introns: 26/3; 82/3; 130/3; 195/3; 241/3; 613/1; 762/1

Query Match 91.7%; Score 22; DB 2; Length 848;
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 Db 630 GAAAPS 635

RESULT 48

T20429
 hypothetical protein E03A3.6 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C;Accession: T20429
 R;Gardiner, A.
 Submitted to the EMBL Data Library, October 1994
 A;Cross references: EMBL:Z29274
 A;Accession: T20429
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-875 <WIL>
 A;Cross references: EMBL:Z28112; PIDN:CAA86231.1; GSPDB:GN00021; CESP:E03A3.6
 A;Experimental source: clone E03A3
 C;Genetics:
 A;Gene: CESP:E03A3.6
 A;Map position: 3
 A;Introns: 58/2; 161/3; 200/1; 257/3; 436/3; 629/3; 735/1; 762/3
 C;Superfamily: *Caenorhabditis elegans* hypothetical protein E03A3.6

Query Match 91.7%; Score 22; DB 2; Length 875;
 Best Local Similarity 66.7%; Pred. No. 5.5e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gtxxps 6
 Db 827 GTSAPS 832

RESULT 46

JS0747
 regulatory protein Spi - rat (Norway rat)
 C;Species: *Rattus norvegicus* (Norway rat)
 C;Date: 0-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
 C;Accession: JS0747; S25287
 R;Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami, submitted to JIPID, September 1992
 A;Reference number: JS0747
 A;Accession: JS0747
 A;Molecule type: mRNA
 A;Residues: 1-788 <IMA>
 A;Cross references: DBJ:DI2768; PIDN:9220911; PID:di1002730; PID:g220912
 R;Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami, submitted to JIPID, September 1992
 A;Title: Two regulatory proteins that bind to the basic transcription element (BPE), a c
 A;Reference number: S25287; MUID:93010958
 A;Accession: S25287

RESULT 49
 A47521
 Capsid protein - giardiaivirus GLV
 C;Species: giardiaivirus, GLV
 C;Date: 21-Jan-1994 #sequence_revision 13-Feb-1998 #text_change 16-Jul-1999
 C;Accession: A47521
 R;Wang, A.L.; Yang, H.M.; Shen, K.A.; Wang, C.C.
 Proc. Natl. Acad. Sci. U.S.A. 90, 8595-8599, 1993
 A;Title: Giardiaivirus double-stranded RNA genome encodes a capsid polypeptide and a gag
 A;Reference number: A47521; MUID:93391401
 A;Contents: host Giardia lamblia
 A;Accession: A47521
 A;Molecule type: genomic RNA
 A;Residues 1-886 <WAN>
 A;Cross-references: GB:L113218; NID:91352866; PIDN:AAB01578.1; PID:gi1352867
 A;Note: sequence modified after extraction from NCBI backbone (NCBIN:13759, NCBIP:13759
 C;Superfamily: giardiaivirus capsid protein
 C;Keywords: Capsid protein

Query Match Score 22; DB 1; Length 886;
 Best Local Similarity 91.7%; Pred. No. 6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 9ttxxps 6
 |||
 Db 796 GTAAPS 801

RESULT 50
 T40290
 Hypothetical protein SPBC354.10 - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T40290
 R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
 Submitted to the EMBL Data Library, March 1998
 A;Reference number: 221918
 A;Accession: T40290
 A;Status: Preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues 1-963 <WOO>
 A;Cross-references: EMBL:AL022071; PIDN:CAA17810.1; GSPDB:GN00067; SPDB:SPBC354.10
 C;Experimental source: strain 972h-; cosmid c354
 A;Gene: SPBC354.10
 A;Map position: 2

Query Match Score 22; DB 2; Length 963;
 Best Local Similarity 91.7%; Pred. No. 6.5e-02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 9ttxxps 6
 |||
 Db 730 GTATPS 735

Search completed: September 3, 2002, 09:25:53
 Job time: 45 sec

AC	Q90477;	DT	15-DEC-1998 (Rel. 37, Created)
		DT	15-DEC-1998 (Rel. 37, Last sequence update)
		DT	15-DEC-1998 (Rel. 37, Last annotation update)
		DE	Myoblast determination protein 1 homolog (Myogenic factor 1).
		GN	MYOD.
		OS	Brachydanio rerio (zebrafish) (zebra danio).
		RA	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
		NCBI_TaxID	955;
		RN	[1]
		RP	SEQUENCE FROM N.A.
		RX	MEDLINE=96152221; PubMed=8565839;
		RA	Neinberg E.S., Allende M.L., Kelly C.S., Abdelhamid A., Murakami T., Andermann P., Doerr O.G., Grunwald D.J., Riggleman B.;
		RT	"Developmental regulation of zebrafish MyoD in wild-type, no tail and spadetail embryos.";
		RL	Development 122:271-280(1996).
		CC	-!- FUNCTION: MYOGENESIS.
		CC	-!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER BHLH PROTEIN.
		CC	-!- SUBCELLULAR LOCATION: Nuclear.
		CC	-!- DEVELOPMENTAL STAGE: EXPRESSION IS SPECIFIC TO THE MESODERM OF THE GASTRULA.
		CC	-!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS. "MYOGENIC FACTORS" SUBFAMILY.
		CC	-----
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).
		CC	-----
		CC	EMBL; 236947; CAA85407; 1. -.
		DR	HSSP; P10085; 1MDY.
		DR	ZDB-GENE-980526-561; myod.
		DR	InterPro; IPR002546; Basic.
		DR	InterPro; IPR003015; HHLH_Myc.
		DR	InterPro; IPR001092; HHLH_dim.
		DR	Pfam; PF01586; Basic; 1.
		DR	Pfam; PF00010; HHLH; 1.
		DR	SMART; SM00520; BASIC; 1.
		DR	PROSITE; PS00038; HELIX-LOOP-HELIX; 1.
		DR	Myogenesis; Differentiation; Developmental protein; Nuclear protein; RNA transcription regulation; DNA-binding.
		KW	BASIC DOMAIN.
		FT	DNA_BIND 84
		FT	DOMAIN 97 136 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
		SEQUENCE	275 AA; 30866 MW; 260091DDA75631C CRC64;
		QY	1 gtxxps 6
		Db	253 GTTAPS 258
		RESULT	4
		MTR7_MOUSE	STANDARD;
		ID	MTR7_MOUSE
		AC	Q9Z2C9;
		DT	16-OCT-2001 (Rel. 40, Created)
		DT	16-OCT-2001 (Rel. 40, Last sequence update)
		DT	16-OCT-2001 (Rel. 40, Last annotation update)
		DE	Myotubularin related protein 7 (EC 3.1.3.48) (Fragment).
		GN	MTR7.
		OS	Mus musculus (Mouse).
		PRT;	312 AA.
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	
		DT	
		DT	
		DT	
		DE	
		GN	
		OS	
		PRT;	
		STANDARD;	
		MTR7_MOUSE	
		ID	
		AC	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TAXID=10090;
[1]
SEQUENCE FROM N.A.
MEDLINE-#9849499; PubMed-#9736772;
Laporte J., Blondeau P., Buij-Belio A., Tentler D., Kretz C., Dahl N.,
Mandel J.-L.;
"Characterization of the myotubularin dual specificity phosphatase
gene family from Yeast to human";
Hum. Mol. Genet. 7:1703-1712 (1998).
-1- FUNCTION: NOT KNOWN, COULD BE A TYROSINE-PHOSPHATASE.
-1- SIMILARITY: BELONGS TO THE MYOTUBULARIN-LIKE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on
use by non-profit institutions as long as its content is in no
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. (See <http://www.isb-sib.ch/announcements.html>
or send an email to license@isb-sib.ch.)

- 1- SIMILARITY: BELONGS TO THE MYOTOBULAR-LIKE FAMILY.
- 2- SIMILARITY: COULD BE A ROSANE PROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

Query Match      91.7%; Score 22; DB 1; Length 379;
Best Local Similarity 66.7%; pred. No. 90;
Matches 4; Conservative 0; Mismatches 2; Indels 0;
                                     gaps 0;
Qy      1 gtxxps 6
Dh      194 GTAGSPS 109

```

RESULT	6
CHS1_LYCES	
ID	CHS1_LYCES
AC	P23418;
DT	01-NOV-1991 (Rel. 20, Created)
DT	01-APR-1993 (Rel. 25, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	Chalcone synthase 1 (EC 2.3.1.74) (Naringenin-chalcone synthase 1).
GN	CHS1
OS	Lycopersicon esculentum (Tomato).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
SPERMATOPHYTA	Magnoliophyta; eudicots; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.	
QC	

Query Match	Best Local Similarity	Score 22;	DB 1;	Length 312;
Matches 4:	Conservative	91.78;	Score 22;	DB 1;
Best Local Similarity	66.78;	Pred. No. 75;	0. Mismatches	0. Cands

RN [1] RP SEQUENCE FROM N.A.
RT TISSUE=Cotyledon, Hypocotyl, and Leaf;
RX MEDLINE=91117196; PubMed=1980324;
RA O'Neill S.D., Tong Y., Soerlein B., Forkmann G., Yoder J.I.;
RE http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=91117196

RESULT	5	T1_TORCA	P19333;	STANDARD;	PRT;	379 AA.
			01-NOV-1990	(Rel. 16, Created)		
			01-NOV-1990	(Rel. 16, Last sequence update)		
			30-MAY-2000	(Rel. 39, Last annotation update)		
				Synaptic vesicle membrane protein VAR-1.		
				Torpedo californica (Pacific electric ray).		
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;		
				Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;		
				Torpediniformes; Torpedinoidae; Torpedinidae; Torpedo.		
						NCBI TaxID:77907

descendant and an ancient derivative molecule. ;
 RL Mol. Gen. Genet. 227:279-288 (1990).
 CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4',2',4',6'-
 CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN.
 CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 COA +
 CC naringenin-chalcone + 3 CO(2).
 CC -1- BIOCHEMISTRY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.
 CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
 CC

for pectoriformes; torpedinoides; torpedinidae; torpedo; torpedo. [1] NCBI TaxID:7787; [1] SEQUENCE FROM N.A. AND PARTIAL SEQUENCE. TISSUE-Electric lobe; MEDLINE=9016593; PubMed=2483112; Linial M., Miller K., Scheiner R.H.; "VAT-1: an abundant membrane protein from Torpedo cholinergic synapses". *Proc Natl Acad Sci USA* 1990; 87: 11211-11215.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

ACR 2000-2001 227 (2001)

-1- FUNCTION: MAY PLAY A CENTRAL ROLE IN THE FUNCTIONS MEDIATED BY SPECIFIC CLASSES OF SYNAPTIC VESICLES.

-1- SUBCELLULAR LOCATION: Membrane.

-1- TISSUE SPECIFICITY: CHOLINERGIC SYNAPTIC VESICLES.

-1- MISCELLANEOUS: SYNTHESIZED IN THE NEURONAL CELL BODIES AND TRANSPORTED TO THE TERMINALS. EACH VESICLE CONTAINS APPROXIMATELY 28 MOLECULES OF VAT-1.

-1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE

SEQUENCE FROM N.A.
STRAIN=CV. RED PONTIAC;
Joung J.-H., Kim H.-S., Choi K.-H., Joung Y.-H., Joung H., Byun S.-M.;
"Characterization of two members of the chalcone synthase gene family from Solanum tuberosum L.;"
(In) Plant Gene Register PGR96-027.
-1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4',2',4",6"-TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE) WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO NARINGENIN.
-1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4 -comaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO(2).
-1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY OF WHICH ARE BRILLYANTLY COLORED.
-1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to licensee@isb-sib.ch).

-! - PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY OF WHICH ARE BRIGHTLY COLORED.

-! - SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

DR EMBL; U47740; AAB67735; 1; .

DR Mendel; 10625; S0LTU;1193;2.

DR InterPro; IPR001099; Chal_stil_synt.

DR Pfam; PF00195; Chal_stil_synt; 1.

DR Pfam; PF02797; Chal_stil_syntC; 1.

DR PROSITE; PS00441; CHALCONE_SYNTH; 1.

KW Flavonoid biosynthesis; Transferase; Acyltransferase; Multigene family.

FT ACT_SITE 164 164 BY SIMILARITY.

SQ SEQUENCE 389 AA; 42548 MW; E7334ALA34DD1CC CRC64;

Query Match 91.7% Score 22; DB 1; Length 389;

Best Local Similarity 66.7% Pred. No. 92;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps

Qy	1 9ttxps 6
Db	23 GTSTPS 28

RESULT 11

CHSJ_PETHY	STANDARD;	PRT;	389 AA.
AC P29298;			
DT 01-AUG-1991 (Rel. 19, Created)			
DT 01-AUG-1993 (Rel. 25, Last sequence update)			
DT 15-JUL-1999 (Rel. 38, Last annotation update)			
DE Chalcone synthase J (EC 2.3.1.74) (Naringenin-chalcone synthase J).			
GN CHSJ.			
OS Petunia hybrida (Petunia).			
OC Buxaceae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyta; eudicots; core eudicots;			
OC Asteridae; euasterids I; Solanales; Solanaceae; Petunia.			
OX NCBI_TAXID=4102;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=CV. VIOLET 30; TISSUE=Leaf;			
RX MEDLINE=90034197; PubMed=2806915;			
RA Koes R.E.; Spelt C.E.; van den Elzen P.J.M.; Mol J.N.M.;			
RT "Cloning and molecular characterization of the chalcone synthase multigene family of Petunia hybrida.",			
RT Gene 81:245-257(1989).			
RL [2]			
RP SEQUENCE OF 71-389 FROM N.A.			
RC STRAIN=CV. WHITE 137; TISSUE=Anther;			
RA van Tunen A.J.;			
RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.			
-! - FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE) WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO NARINGENIN.			
CC -! - CATALYTIC ACTIVITY: 3 malonyl-CoA + 4 -coumaroyl-CoA = CoA + naringenin-chalcone + 3 CO(2).			
CC -! - PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY OF WHICH ARE BRIGHTLY COLORED.			
CC -! - SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

InterPro: IPR01099; chal_stil_synt.
 Pfam: PF00155; Chal_stil_synt; 1.
 Pfam: PF02797; Chal_stil_synt; 1.
 PROSITE: PS00441; CHALCONE_SYNTH; 1.
 Flavonoid biosynthesis; Transferase; Acyltransferase.
 ACT_SITE 164 164 BY SIMILARITY.
 SEQUENCE 389 AA: 42492 MW: B350742DEA19186 CRC64;

```

  LnuL, a1322; Chal33.1; .
  PIR; JS0309; SYPCJU.
  InterPro; IPR01099; Chal_still_synt.
  R InterPro; IPR01099; Chal_still_synt.
  R Pfam; PF00195; Chal_still_synt; 1.
  R Pfam; PF02797; Chal_still_syntC; 1.
  PROSITE; PS00441; CHALCONE_SYNTH; 1.
  R Flavonoid biosynthesis; Transferase; Acyltransferase;
  W Multigene family.
  ACT_SITE 164 BY SIMILARITY.
  T 75 E -> V (IN REF. 2).
  T CONFLICT 75 MW: F2B3CDD82E6FDE7 CRC64:
  SEQUENCE 389 AA: 42558
  O RESULT 13

```

CHS_PHAU	LD	PRF;	389 AA.
P49440;	AC		
01-FEB-1996 (Rel. 33, Created)	DT		
01-FEB-1996 (Rel. 33, Last annotation update)	DT		
15-JUL-1999 (Rel. 38, Last annotation update)	DT		
Chalcone synthase 17 (EC 2.3.1.74) (Naringenin-chalcone synthase 17).	DE		
CHS17.	GN		
Phaseolus vulgaris (Kidney bean) (French bean).	OS		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; tracheophyta;	OC		
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	OC		
eurosid I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.	OC		
NCBI TaxID:3885;	OX		

RESULT 12
HSY_CATRO
D CHSY_CATRO
C 092454;
C 16-OCT-2001 (Ref)
C 40 Created
STANDARD; PRT; 389 AA.

16-Oct-2001 (Rel. 40, Last sequence update)
 16-Oct-2001 (Rel. 40, Last annotation update)
 Chalcone synthase (EC 2.3.1.74) (Naringenin-chalcone synthase).
 CHS.

Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle),
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euanthids I; Gentianales; Apocynaceae; Rauvolfioideae;
Solanaceae; Catharanthus; Catharanthus roseus (L.) G. Don; Catharanthus
RT the Plant defense enzyme chalcone synthase in *Phaseolus vulgaris*.ⁿ
RL Mol. Gen. Genet. 210:219-233 (1987).
-1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4',2',4'-6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC

CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN.
 CC -!- CATALYTIC ACTIVITY : 3 malonyl-CoA + 4 -coumaroyl-CoA = 4 CoA +
 CC
 NCBI_TAXID=4058;
 [1]_X
 N

CC naringenin-chalcone + 3 CO₂.

CC -1: PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.

Plants"; Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4',2',4',6'-TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)

WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO NARINGENIN.

-1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO₂(2).

-1- PATHWAY: PART OF THE BIOTSYNTHETIC PATHWAY FOR ALL CLASSES OF

PLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY OF WHICH ARE BRIGHTLY COLORED.
-1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way used for commercial purposes.

entities require a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

GN	Daucus	Carota (Carrot).
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicots; core eudicots;	
OC	Asteridae; euasterids II; Apiales; Apiaceae; Daucus.	
OX	NCBI_TaxID=4039;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Hirner A. A.; Seitz H.U.;	
RT	"Cloning of two different chalcone synthase isoforms from Daucus carota L. and their differential expression in organs from European wild carrot and in UV-irradiated cell cultures."	
RL	Submitted (JUN-1988) to the EMBL/Genbank/DBJ databases.	
RI	-1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4',2',4'',6''-TETRAHYDROCHALCONE (ALSO TERMED NARINGENIN CHALCONE OR CHALCONE) WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO NARINGENIN.	
CC	-1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4 COA + naringenin-chalcone + 3 CO(2).	
CC	-1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES. MANY OF WHICH ARE BRIGITELY COLORED.	
CC	-1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcements or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; At006780; CAA07245.1;	
DR	Mendel; 36898; Dauga;1193;36898.	
DR	InterPro; IPR001099; Chal_stil_synt.	
DR	Pfam; PF00195; Chal_stil_synt; 1.	
DR	Pfam; PF02797; Chal_stil_synt; 1.	
DR	PROSITE; PS00441; CHALCONE_SYNTH; 1.	
DR	Flavonoid biosynthesis; Transferase; Acyltransferase;	
KW	Multigene family.	
FT	ACT SITE 168	
FT	SEQUENCE 397 AA; 43559 MW; 9386F44B9132EDEE CRC64;	
SQ	-----	
Query Match	91.7%; Score 22; DB 1; Length 397;	
Best Local Similarity	66.7%; Pred. No. 94;	
Matches 4;	Conservative 0; Mismatches 2; Indels 0; Gaps	
Qy	1 gttxps 6	
	27 GTATPS 32	
Db		
RESULT	17	
CHSY_CALCH	STANDARD;	
ID	CHSY_CALCH	
PRN	398 AA.	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=L 01; TISSUE=Petal;	
RA	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicots; core eudicots;	
OC	Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;	
OC	Callistephus chinensis (China aster).	
OX	NCBI_TaxID=13379;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=L 01; TISSUE=Petal;	
RA	Eukaryota M.; Sommer H.; Forkmann G.; Wassenegger M.; Sommer H.; Forkmann G.; J. Wassenegger M.; Sommer H.; Forkmann G.; Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.	
RL		

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: V01538; CA324779.1; -.

DR PTR: SA2523; SA4223.

DR InterPro; IPR01099; Chal_stil_synt.

DR Pfam; PF00195; Chal_stil_synt; 1.

DR Pfam; PF02797; Chal_stil_syntC; 1.

DR PROSITE; PS00441; CHALCONE_SYNTH; 1.

KW Flavonoid biosynthesis; Transferase; Acyltransferase.

FT ACT SITE 169 169 BY SIMILARITY.

SEQUENCE 398 AA; 43735 MW; E8B19149AD3DAB1 CRC64;

Query Match 91.7%; Score 22; DB 1; Length 398;

Best Local Similarity 66.7%; Pred. No. 94;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6

Db 28 GTATPS 33

RESULT 19

CHSD_PETHY STANDARD; PRT; 419 AA.

ID CHSD_PETHY

AC P22925;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

GN Chalcone synthase D (BC 2.3.1.74) (Naringenin-chalcone synthase D).

OS Petunia hybrida (Petunia).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Petunia.

NCBI_TAXID=4102;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. VIOLET 30; TISSUE=Leaf;

RX MEDLINEP=90034197; PubMed=2006915;

RA Koes R.E., Speit C.E., van den Elzen P.J.M., Mol J.N.M.;

RT "Cloning and molecular characterization of the chalcone synthase multigene family of Petunia hybrida.";

RL Gene 81:245-257 (1989).

-1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4',2',4',6'-TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE) WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO NARINGENTIN.

CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 COA + naringenin-chalcone + 3 CO(2).

CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF FLAVONOLS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY OF WHICH ARE BRIGHTLY COLORED.

CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL: X14593; CAA32733.1; -.

CC DR PIR; JSD312; SYP0CD.

CC DR InterPro; IPR01099; Chal_stil_synt.

CC DR Pfam; PF00195; Chal_stil_synt; 1.

CC DR Pfam; PF02797; Chal_stil_syntC; 1.

CC DR PROSITE; PS00441; CHALCONE_SYNTH; 1.

CC KW Flavonoid biosynthesis; Transferase; Acyltransferase; Multigene family.

FT	ACT_SITE	164	BY SIMILARITY.
FT	SEQUENCE	419 AA;	45979 MW;
FT	SEQUENCE	55827338DC812ACD	CRC64;
Query	1 gtxxps 6		
Db	1	Score 22; Pred. No. 99; Mismatches 0;	Length 4 2; Indel 2;
Db	23 GSTATPS 28		
RESULT 20			
UGTB_CAEEL	STANDARD;	PRT;	435 AA.
ID UGTB_CAEEL			
AC Q22180;			
DT 01-NOV-1997 (Rel. 35, Created)			
DT 01-NOV-1997 (Rel. 35, Last sequence update)			
DT 16-OCT-2001 (Rel. 40, Last annotation update)			
DE Putative UDP-galucuronosyltransferase UGT11 (EC 2.4.1.17)			
GN UGT11 OR TOH1.7.			
OS Caenorhabditis elegans.			
CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;			
CC Rhabditidae; Peledorinae; Caenorhabditis.			
CX NCBI_TAXID=6239;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA STRAIN-BRISTOL N2;			
RA Harris P.			
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases			
CC -!- CATALYTIC ACTIVITY: UDP-glucurononate + acceptor = UDP-beta-D-glucurononoside.			
CC -!- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE			
CC			
CC This SWISS-PROT entry is copyright. It is produced thrc			
CC between the Swiss Institute of Bioinformatics and the			
CC European Bioinformatics Institute. There are no r			
CC use by non-profit institutions as long as its conte			
CC modified and this statement is not removed. Usage by			
CC entities requires a license agreement (See <a href="http://www.i			
CC or send an email to license@ibis.sib.ch).			
CC			
DR EMBL; Z78200; CAB01584.1;			
DR WormPep; T04H1.7; CE13179.			
DR InterPro; IPR002213; UDPGT.			
DR Pfam; PF00201; UDPGT; 1.			
DR PROSTRE; PS00375; UDPGT; 1.			
DR HYPOTHETICAL PROTEIN; TRANSFERASE; Glycosyltransferase;			
KW Multiceline family.			
SEQUENCE 435 AA;	50020 MW;	AF04A0DE7208CBF3	CRC64;
SQ			
Query	1 gtxxps 6		
Db	1	Score 22; Pred. No. 1e+02; Mismatches 0;	Length 4 2; Indel 2;
Db	300 GTTTPS 305		
RESULT 21			
DNAK_PAVLU	STANDARD;	PRT;	629 AA.
ID DNAK_PAVLU			
AC P30722;			
DT 01-APR-1993 (Rel. 25, Created)			
DT 01-APR-1993 (Rel. 25, Last sequence update)			
DT 16-OCT-2001 (Rel. 40, Last annotation update)			
DE Chaperone protein dnaK (heat shock protein 70) (heat sh			
DE protein) (HSP70).			
DN DNAK_PAVLU (Monochrysis lutheri).			
OS Pavlova lutheri (Monochrysis lutheri).			

OG Chloroplast.
OC Eukaryota; Haptophyceae; Pavlovales; Pavlova.
OX NCBI_TaxID:2832;
RN [1] SEQUENCE FROM N.A.
RX MEDLINE-92163012; PubMed=1536924;
RA Scaramuzzi C.D.; Stokes H.W.; Hiller R.G.;
RT "Heat shock Hsp70 protein is chloroplast-encoded in the chromophytic
RT alga *Lutheria*."
RL Plant Mol. Biol. 18:467-476(1992).
CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59555; CAAK2154.1; -.
DR PIR; S20516; S20516.
DR HSSP; P04475; 1DG4.
DR InterPro; IPR001023; HSP70.
DR Pfam; PF00012; HSP70.
DR PRINTS; PR00301; HEATSHOCK70.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR Chaperone; ATP-binding; Heat shock; Chloroplast.
KW SEQUENCE 629 AA; 68792 MW; C3880DC369979D56 CRC64;
SQ -----
Query Match 22 Score 22; DB 1; Length 629;
Best Local Similarity 66.7%; Pred. No. 1.e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DE Transcript factor Sp1.
DE Sp1 OR TSFPI.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1] SEQUENCE OF 4-785 FROM N.A.
RX TISSUE-Cervical carcinoma.
RA Haggart M.H.; Laurner A.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN [2] SEQUENCE OF 1-558 FROM N.A.
RX Takahara T.; Kanzu S.; Yanagisawa S.; Akanuma H.;
RT "Heterogeneous Sp1 mRNAs in human HepG2 cells include a product of
RT homotypic trans-splicing.";
RL J. Biol. Chem. 275:38067-38072(2000).
RN [3] SEQUENCE OF 90-785 FROM N.A. AND SEQUENCE OF 359-375 AND 670-675.
RX MEDLINE-68080466; PubMed=3319186;
RA Kadonaga J.T.; Carner K.R.; Masiarz F.R.; Tjian R.;
RT "Isolation of cDNA encoding transcription factor Sp1 and functional

RT analysis of the DNA binding domain.";
RL Cell 51:1079-1090(1988).
RN [4] O GLYCOSYLATION.
RP MEDLINE-8903041; PubMed=3139301;
RX MEDLINE-9728212; PubMed=9065444;
RA Jackson S.P.; Tjian R.;
RT "O-glycosylation of eukaryotic transcription factors: implications
RT for mechanisms of transcriptional regulation.";
RL Cell 55:125-133(1988).
RN [5] RSTRUCTURE BY NMR OF 654-684 AND 684-712.
RX MEDLINE-9728212; PubMed=9065444;
RA Narayan V.A.; Kriwacki R.W.; Caradonna J.P.;
RT "Structures of zinc finger domains from transcription factor Sp1.
RT Insights into sequence-specific protein-DNA recognition.";
RL J. Biol. Chem. 272:7801-7809(1997).
RN [6] IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES.
RP MEDLINE=96224025; PubMed=8626723;
RX Parks C.L.; Shenk T.;
RA "The serotonin 1a receptor gene contains a TATA-less promoter that
RT responds to MAZ and Sp1.";
RL J. Biol. Chem. 271:4417-4430(1996).
CC -1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES mRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES. CAN INTERACT WITH G/C-RICH MOTIFS FROM
CC SEROTONIN RECEPTOR PROMOTER.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PIM- O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS.
CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AF252284; AAF67726.1; -.
CC DR EMBL; AB039486; BAB13476.1; -.
CC DR J03133; AAA61154.1; -.
CC DR J29635; A29635.
CC DR PDB; 1SP1; 21-APR-97.
CC DR PDB; 1SP2; 21-APR-97.
CC DR TRANSFAC; T00759; -.
CC DR Glycositedb; P08047; -.
CC DR MM; 189906; -.
CC DR InterPro; IPR000822; Znf-C2H2.
CC DR Pfam; PF00056; zf-C2H2; 3.
CC DR SMART; SM00355; Znf_C2H2; 3.
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
CC DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
CC KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
DNA-binding; Nuclear protein; Repeat; Glycoprotein; 3D-structure.
FT DOMAIN 626 708
FT 2N_FING 626 656
FT 2N_FING 656 680
FT 2N_FING 686 708
FT CONFLICT 366 366
FT CONFLICT 670 670
SQ SEQUENCE 785 AA; 80693 MW; 43893DBF6518B9EA CRC64;
ZINC FINGERS.
Query Match 91.7%; Score 22; DB 1; Length 785;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DT	15-JUL-1999	(Rel. 38, Created)
DT	15-JUL-1999	(Rel. 38, Last sequence update)
DT	16-OCT-2001	(Rel. 40, Last annotation update)
DE	Serine/threonine-protein kinase ULK1 (EC 2.7.1.-)	(Unc-51-like kinase 1).
GN	ULK1	
GN	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TAXID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	SEQUENCE FROM N.A.	
RX	SEQUENCE FROM N.A.	
RC	TISSUE=Liver;	
RX	TISSUE=Liver;	
RA	Medline=93010958; PubMed=1356762;	
RA	Imataki H., Sogawa K., Yasumoto K., Kikuchi Y., Sesano K.,	
RA	Yamamoto A., Hayami M., Fujii-Kuriyama Y.;	
RT	"Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene,"	
RT	EMBO J. 11:3663-3671(1992).	
CC	-1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY ACTIVATES mRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL RECOGNITION SITES.	
CC	-1- SUBCELLULAR LOCATION: Nuclear. That bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene.	
CC	-1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS (BY SIMILARITY)	
CC	-1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	CC	This SWISS-PROT entry is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	EMBL; AF04558; AAC32326_1; -.	
DR	EMBL; P00523; 2PTK.	
DR	MM; 633168; -.	
DR	PROSITE; PS000719; Euk_pk kinase.	
DR	InterPro; IPR002290; Ser_thr_pk kinase.	
DR	InterPro; IPR001245; Tyr_pk kinase.	
DR	Pfam; PF000659; kinase_1.	
DR	PRINTS; PRO00109; TYRKINASE.	
DR	SMART; SM00220; S_TK_C; 1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
DR	PROSITE; PS55011; PROTEIN_KINASE_DOM; 1.	
KW	Transferase; Serine/threonine-protein kinase; ATP-binding.	
FT	DOMAIN_16_278	
FT	NF_BIND_22_30	
FT	BINDING_46_46	
FT	ACT_SITE_138_138	
FT	DOMAIN_297_310	
SQ	SEQUENCE_1050_AA; 112601 MW; 4ED9B94028E3C138 CRC64;	
Qy	1 gtxxps 6	
Db	745 GRSSPS 750	
RESULT	25	
GNRP_MOUSE	STANDARD;	
AC	P27671;	
DT	01-AUG-1992 (Rel. 23, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Guanine nucleotide releasing protein (GNRP) (RAS-specific nucleotide exchange factor GNRP, OR GNRC25 OR GRP).	
GN	GNRP, OR GNRC25, OR GRP.	
RESULT	24	
ULK1_HUMAN	STANDARD;	
AC	U0K1_HUMAN	
AC	Q53855;	
Qy	1 gtxxps 6	
Db	739 GTATPS 744	
RESULT	23	
SP1_RAT	STANDARD;	
ID	SP1_RAT	
AC	Q01714;	
DT	01-APR-1993 (Rel. 25, Created)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Transcription factor Sp1.	
GN	SP1.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
ON	NCBI_TAXID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	SEQUENCE FROM N.A.	
RA	Medline=93010958; PubMed=1356762;	
RA	Imataki H., Sogawa K., Yasumoto K., Kikuchi Y., Sesano K., Kobayashi A., Hayami M., Fujii-Kuriyama Y.;	
RT	"Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene,"	
RT	EMBO J. 11:3663-3671(1992).	
CC	-1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY ACTIVATES mRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL RECOGNITION SITES.	
CC	-1- SUBCELLULAR LOCATION: Nuclear. That bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene.	
CC	-1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS (BY SIMILARITY)	
CC	-1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	EMBL; D12768; BAA0225_1; -.	
DR	PIR; JS0747; JS0747.	
DR	HSSP; P08047; 1SP1.	
DR	TRANSFAC; T00754; -.	
DR	InterPro; IPR000822; Znf-C2H2.	
DR	InterPro; IPR000822; Znf-C2H2.	
DR	PRINTS; PRO0048; zf-C2H2; 3.	
DR	SMART; SM00355; Znf_C2H2; 3.	
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.	
DR	PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.	
KW	Transcription regulation; Activator; Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat; Glycoprotein.	
KW	ZINC FINGERS.	
FT	DOMAIN_629_711	
FT	ZNFING_629_653	
FT	ZNFING_659_683	
FT	ZNFING_689_711	
SQ	SEQUENCE_788_AA; 81015 MW; AA2B0CAB81AAABB80C CRC64;	
Qy	1 gtxxps 6	
Db	739 GTATPS 744	
RESULT	23	
GNRP_MOUSE	STANDARD;	
AC	P27671;	
DT	01-AUG-1992 (Rel. 23, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Guanine nucleotide releasing protein (GNRP) (RAS-specific nucleotide exchange factor GNRP, OR GNRC25 OR GRP).	
GN	GNRP, OR GNRC25, OR GRP.	
RESULT	22	
SP1_RAT	STANDARD;	
ID	SP1_RAT	
AC	Q01714;	
DT	01-APR-1993 (Rel. 25, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Transcription factor Sp1.	
GN	SP1.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
ON	NCBI_TAXID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	SEQUENCE FROM N.A.	
RA	Medline=93010958; PubMed=1356762;	
RA	Imataki H., Sogawa K., Yasumoto K., Kikuchi Y., Sesano K., Kobayashi A., Hayami M., Fujii-Kuriyama Y.;	
RT	"Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene,"	
RT	EMBO J. 11:3663-3671(1992).	
CC	-1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY ACTIVATES mRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL RECOGNITION SITES.	
CC	-1- SUBCELLULAR LOCATION: Nuclear. That bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene.	
CC	-1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS (BY SIMILARITY)	
CC	-1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	EMBL; AF04558; AAC32326_1; -.	
DR	EMBL; P00523; 2PTK.	
DR	MM; 633168; -.	
DR	PROSITE; PS000719; Euk_pk kinase.	
DR	InterPro; IPR002290; Ser_thr_pk kinase.	
DR	InterPro; IPR001245; Tyr_pk kinase.	
DR	Pfam; PF000659; kinase_1.	
DR	PRINTS; PRO00109; TYRKINASE.	
DR	SMART; SM00220; S_TK_C; 1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
DR	PROSITE; PS55011; PROTEIN_KINASE_DOM; 1.	
KW	Transferase; Serine/threonine-protein kinase; ATP-binding.	
FT	DOMAIN_16_278	
FT	NF_BIND_22_30	
FT	BINDING_46_46	
FT	ACT_SITE_138_138	
FT	DOMAIN_297_310	
SQ	SEQUENCE_1050_AA; 112601 MW; 4ED9B94028E3C138 CRC64;	
Qy	1 gtxxps 6	
Db	745 GRSSPS 750	
RESULT	25	
GNRP_MOUSE	STANDARD;	
AC	P27671;	
DT	01-AUG-1992 (Rel. 23, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Guanine nucleotide releasing protein (GNRP) (RAS-specific nucleotide exchange factor GNRP, OR GNRC25 OR GRP).	
GN	GNRP, OR GNRC25, OR GRP.	
RESULT	24	
ULK1_HUMAN	STANDARD;	
ID	ULK1_HUMAN	
AC	Q53855;	
Qy	1 gtxxps 6	
Db	739 GTATPS 744	
RESULT	23	
SP1_RAT	STANDARD;	
ID	SP1_RAT	
AC	Q01714;	
DT	01-APR-1993 (Rel. 25, Created)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Transcription factor Sp1.	
GN	SP1.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
ON	NCBI_TAXID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	SEQUENCE FROM N.A.	
RA	Medline=93010958; PubMed=1356762;	
RA	Imataki H., Sogawa K., Yasumoto K., Kikuchi Y., Sesano K., Kobayashi A., Hayami M., Fujii-Kuriyama Y.;	
RT	"Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene,"	
RT	EMBO J. 11:3663-3671(1992).	
CC	-1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY ACTIVATES mRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL RECOGNITION SITES.	
CC	-1- SUBCELLULAR LOCATION: Nuclear. That bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene.	
CC	-1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS (BY SIMILARITY)	
CC	-1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	EMBL; AF04558; AAC32326_1; -.	
DR	EMBL; P00523; 2PTK.	
DR	MM; 633168; -.	
DR	PROSITE; PS000719; Euk_pk kinase.	
DR	InterPro; IPR002290; Ser_thr_pk kinase.	
DR	InterPro; IPR001245; Tyr_pk kinase.	
DR	Pfam; PF000659; kinase_1.	
DR	PRINTS; PRO00109; TYRKINASE.	
DR	SMART; SM00220; S_TK_C; 1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
DR	PROSITE; PS55011; PROTEIN_KINASE_DOM; 1.	
KW	Transferase; Serine/threonine-protein kinase; ATP-binding.	
FT	DOMAIN_16_278	
FT	NF_BIND_22_30	
FT	BINDING_46_46	
FT	ACT_SITE_138_138	
FT	DOMAIN_297_310	
SQ	SEQUENCE_1050_AA; 112601 MW; 4ED9B94028E3C138 CRC64;	
Qy	1 gtxxps 6	
Db	745 GRSSPS 750	
RESULT	25	
GNRP_MOUSE	STANDARD;	
AC	P27671;	
DT	01-AUG-1992 (Rel. 23, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Guanine nucleotide releasing protein (GNRP) (RAS-specific nucleotide exchange factor GNRP, OR GNRC25 OR GRP).	
GN	GNRP, OR GNRC25, OR GRP.	
RESULT	24	
ULK1_HUMAN	STANDARD;	
ID	ULK1_HUMAN	
AC	Q53855;	
Qy	1 gtxxps 6	
Db	739 GTATPS 744	
RESULT	23	
SP1_RAT	STANDARD;	
ID	SP1_RAT	
AC	Q01714;	
DT	01-APR-1993 (Rel. 25, Created)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Transcription factor Sp1.	
GN	SP1.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
ON	NCBI_TAXID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	SEQUENCE FROM N.A.	
RA	Medline=93010958; PubMed=1356762;	
RA	Imataki H., Sogawa K., Yasumoto K., Kikuchi Y., Sesano K., Kobayashi A., Hayami M., Fujii-Kuriyama Y.;	
RT	"Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene,"	
RT	EMBO J. 11:3663-3671(1992).	
CC	-1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY ACTIVATES mRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL RECOGNITION SITES.	
CC	-1- SUBCELLULAR LOCATION: Nuclear. That bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat p-450 _{A1} gene.	
CC	-1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS (BY SIMILARITY)	
CC	-1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	EMBL; AF04558; AAC32326_1; -.	
DR	EMBL; P00523; 2PTK.	
DR	MM; 633168; -.	
DR	PROSITE; PS000719; Euk_pk kinase.	
DR	InterPro; IPR002290; Ser_thr_pk kinase.	
DR	InterPro; IPR001245; Tyr_pk kinase.	
DR	Pfam; PF000659; kinase_1.	
DR	PRINTS; PRO00109; TYRKINASE.	
DR	SMART; SM00220; S_TK_C; 1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
DR	PROSITE; PS55011; PROTEIN_KINASE_DOM; 1.	
KW	Transferase; Serine/threonine-protein kinase; ATP-binding.	
FT	DOMAIN_16_278	
FT	NF_BIND_22_30	
FT	BINDING_46_46	
FT	ACT_SITE_138_138	
FT	DOMAIN_297_310	
SQ	SEQUENCE_1050_AA; 112601 MW; 4ED9B94028E3C138 CRC64;	
Qy	1 gtxxps 6	
Db	745 GRSSPS 750	
RESULT	25	
GNRP_MOUSE	STANDARD;	
AC	P27671;	
DT	01-AUG-1992 (Rel. 23, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Guanine nucleotide releasing protein (GNRP) (RAS-specific nucleotide exchange factor GNRP, OR GNRC25 OR GRP).	
GN	GNRP, OR GNRC25, OR GRP.	

Mus musculus (Mouse). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathini; Muridae; Murinae; Mus. [1] SEQUENCE FROM N.A. STRAIN=BALB/C; MEDLINE=93010996; PubMed=1396590; Cen H.; Lowy D.D.; "Isolation of multiple mouse cDNAs with coding homology to Saccharomyces cerevisiae CDC25: identification of a region related to the Yav, Dbl and CDC4."; EMBO J. 11:4007-4015 (1992). [2] SEQUENCE OF 791-1262 FROM N.A. STRAIN=SWISS; TISSUE=brain; MEDLINE=92289680; PubMed=1376246; Martegani E., Vahoni M., Zippel R., Cocchetti P., Brambilla R., Ferrari C., Sturani E.P., Alberghina L.; "Cloning by functional complementation of a mouse cDNA encoding a homologue of CDC25 (a Saccharomyces cerevisiae RAS activator)."; EMBO J. 11:2151-2157(1992). [3] SEQUENCE OF 1031-1226 FROM N.A. MEDLINE=92357779; PubMed=1379731; Wei W., Mosteller R.D., Sanyal P., Gonzales E., McKinney D., Dasgupta C., Li P., Liu B.X., Broek D., "Identification of a mammalian gene structurally and functionally related to the CDC25 gene of <i>Saccharomyces cerevisiae</i> ."; Proc. Natl. Acad. Sci. U.S.A. 89:7100-7104 (1992). -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP. -1- TISSUE SPECIFICITY: BRAIN. -1- SIMILARITY: CONTAINS 2 PH DOMAINS. -1- SIMILARITY: CONTAINS 1 DBL HOMOLOGY DOMAIN (DH). -1- SIMILARITY: CONTAINS 1 RASGEF DOMAIN. -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.	<p>FT DOMAIN 208 233 IQ. FT DOMAIN 244 430 DH. FT DOMAIN 460 588 PH 2. FT DOMAIN 1025 1259 RASGEF. FT CONFLICT 1033 1033 E->D (IN REF. 3). SQ SEQUENCE 1262 AA; 144101 MW; 38BF68F/C228DC8 CRC64;</p> <p>Query Match 91.7%; Score 22; DB 1; Length 1262; Best Local Similarity 66.7%; Pred. No. 2 8e+02; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;</p> <p>Qy 1 qtxxps 6 Db 734 Gtssps 739</p> <p>RESULT 26 CID_DRONE STANDARD; PRT; 1377 AA. ID CID_DRONE AC P19538; DT 01-NOV-1990 (Rel. 16, Created) DT 01-NOV-1990 (Rel. 16, Last sequence update) DT 01-FEB-1996 (Rel. 33, Last annotation update) CID GN Drosophila interrupta (Fruit fly). OS Drosophila; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidea; Drosophilidae; Drosophila. NCBI_TAXID=7227; RN [1] RP SEQUENCE FROM N.A. RC STRAIN=OREGON-R; RX MEDLINE=90346286; PubMed=2166702; RA Orenic T.V.; Slusarski D.C.; Kroll K.L.; Holmgren R.A.; RT "Cloning and characterization of the segment polarity gene <i>cubitus lateralis</i> in <i>Drosophila</i>."; RN [2] RP SEQUENCE FROM N.A. RC STRAIN=OREGON-R; RX MEDLINE=92146935; PubMed=1686006; RA Berry A.J.; Ajijo J.W.; Kretzman M.; RT "Lack of polymorphism on the <i>Drosophila</i> fourth chromosome resulting from selection"; RN Genetics 129:111-117 (1991). RL PIR; S20730; S20730. PIR; S22693; S22693. MGI; MGI:9694; Rasgrf1. InterPro; IPR001331; GDS_CDC24. InterPro; IPR000048; IQ. InterPro; IPR011849; PH. InterPro; IPR000651; RasGDP. InterPro; IPR001895; RasGDP. InterPro; IPR000219; RhoGEF. Pfam; PF00612; IQ; -. Pfam; PF00616; PH; 2. Pfam; PF00617; RasGEF; 1. Pfam; PF00618; RasGEF; 1. Pfam; PF00621; RhoGEF; 1. SMART; SMART0015; IQ; 1. SMART; SMART0233; PH; 2. SMART; SMART0147; RhoGEF; 1. SMART; SMART0229; RasGEF; 1. SMART; SMART0325; RhoGEF; 1. PROSITE; PS55010; DH; 2; 1. PROSITE; PS50741; DH; 1; 1. PROSITE; PS50720; GDS_CDC25; 1. PROSITE; PS550096; IQ; 1. PROSITE; PS55003; PH_DOMAIN; 2. Guanine-nucleotide releasing factor; Repeat. DOMA1; 22; 10. DOMA2; 22; 10. DOMA3; 22; 10. DOMA4; 22; 10. DOMA5; 22; 10. DOMA6; 22; 10. DOMA7; 22; 10. DOMA8; 22; 10. DOMA9; 22; 10. DOMA10; 22; 10. DOMA11; 22; 10. DOMA12; 22; 10. DOMA13; 22; 10. DOMA14; 22; 10. DOMA15; 22; 10. DOMA16; 22; 10. DOMA17; 22; 10. DOMA18; 22; 10. DOMA19; 22; 10. DOMA20; 22; 10. DOMA21; 22; 10. DOMA22; 22; 10. DOMA23; 22; 10. DOMA24; 22; 10. DOMA25; 22; 10. DOMA26; 22; 10. DOMA27; 22; 10. DOMA28; 22; 10. DOMA29; 22; 10. DOMA30; 22; 10. DOMA31; 22; 10. DOMA32; 22; 10. DOMA33; 22; 10. DOMA34; 22; 10. DOMA35; 22; 10. DOMA36; 22; 10. DOMA37; 22; 10. DOMA38; 22; 10. DOMA39; 22; 10. DOMA40; 22; 10. DOMA41; 22; 10. DOMA42; 22; 10. DOMA43; 22; 10. DOMA44; 22; 10. DOMA45; 22; 10. DOMA46; 22; 10. DOMA47; 22; 10. DOMA48; 22; 10. DOMA49; 22; 10. DOMA50; 22; 10. DOMA51; 22; 10. DOMA52; 22; 10. DOMA53; 22; 10. DOMA54; 22; 10. DOMA55; 22; 10. DOMA56; 22; 10. DOMA57; 22; 10. DOMA58; 22; 10. DOMA59; 22; 10. DOMA60; 22; 10. DOMA61; 22; 10. DOMA62; 22; 10. DOMA63; 22; 10. DOMA64; 22; 10. DOMA65; 22; 10. DOMA66; 22; 10. DOMA67; 22; 10. DOMA68; 22; 10. DOMA69; 22; 10. DOMA70; 22; 10. DOMA71; 22; 10. DOMA72; 22; 10. DOMA73; 22; 10. DOMA74; 22; 10. DOMA75; 22; 10. DOMA76; 22; 10. DOMA77; 22; 10. DOMA78; 22; 10. DOMA79; 22; 10. DOMA80; 22; 10. DOMA81; 22; 10. DOMA82; 22; 10. DOMA83; 22; 10. DOMA84; 22; 10. DOMA85; 22; 10. DOMA86; 22; 10. DOMA87; 22; 10. DOMA88; 22; 10. DOMA89; 22; 10. DOMA90; 22; 10. DOMA91; 22; 10. DOMA92; 22; 10. DOMA93; 22; 10. DOMA94; 22; 10. DOMA95; 22; 10. DOMA96; 22; 10. DOMA97; 22; 10. DOMA98; 22; 10. DOMA99; 22; 10. DOMA100; 22; 10. DOMA101; 22; 10. DOMA102; 22; 10. DOMA103; 22; 10. DOMA104; 22; 10. DOMA105; 22; 10. DOMA106; 22; 10. DOMA107; 22; 10. DOMA108; 22; 10. DOMA109; 22; 10. DOMA110; 22; 10. DOMA111; 22; 10. DOMA112; 22; 10. DOMA113; 22; 10. DOMA114; 22; 10. DOMA115; 22; 10. DOMA116; 22; 10. DOMA117; 22; 10. DOMA118; 22; 10. DOMA119; 22; 10. DOMA120; 22; 10. DOMA121; 22; 10. DOMA122; 22; 10. DOMA123; 22; 10. DOMA124; 22; 10. DOMA125; 22; 10. DOMA126; 22; 10. DOMA127; 22; 10. DOMA128; 22; 10. DOMA129; 22; 10. DOMA130; 22; 10. DOMA131; 22; 10. DOMA132; 22; 10. DOMA133; 22; 10. DOMA134; 22; 10. DOMA135; 22; 10. DOMA136; 22; 10. DOMA137; 22; 10. DOMA138; 22; 10. DOMA139; 22; 10. DOMA140; 22; 10. DOMA141; 22; 10. DOMA142; 22; 10. DOMA143; 22; 10. DOMA144; 22; 10. DOMA145; 22; 10. DOMA146; 22; 10. DOMA147; 22; 10. DOMA148; 22; 10. DOMA149; 22; 10. DOMA150; 22; 10. DOMA151; 22; 10. DOMA152; 22; 10. DOMA153; 22; 10. DOMA154; 22; 10. DOMA155; 22; 10. DOMA156; 22; 10. DOMA157; 22; 10. DOMA158; 22; 10. DOMA159; 22; 10. DOMA160; 22; 10. DOMA161; 22; 10. DOMA162; 22; 10. DOMA163; 22; 10. DOMA164; 22; 10. DOMA165; 22; 10. DOMA166; 22; 10. DOMA167; 22; 10. DOMA168; 22; 10. DOMA169; 22; 10. DOMA170; 22; 10. DOMA171; 22; 10. DOMA172; 22; 10. DOMA173; 22; 10. DOMA174; 22; 10. DOMA175; 22; 10. DOMA176; 22; 10. DOMA177; 22; 10. DOMA178; 22; 10. DOMA179; 22; 10. DOMA180; 22; 10. DOMA181; 22; 10. DOMA182; 22; 10. DOMA183; 22; 10. DOMA184; 22; 10. DOMA185; 22; 10. DOMA186; 22; 10. DOMA187; 22; 10. DOMA188; 22; 10. DOMA189; 22; 10. DOMA190; 22; 10. DOMA191; 22; 10. DOMA192; 22; 10. DOMA193; 22; 10. DOMA194; 22; 10. DOMA195; 22; 10. DOMA196; 22; 10. DOMA197; 22; 10. DOMA198; 22; 10. DOMA199; 22; 10. DOMA200; 22; 10. DOMA201; 22; 10. DOMA202; 22; 10. DOMA203; 22; 10. DOMA204; 22; 10. DOMA205; 22; 10. DOMA206; 22; 10. DOMA207; 22; 10. DOMA208; 22; 10. DOMA209; 22; 10. DOMA210; 22; 10. DOMA211; 22; 10. DOMA212; 22; 10. DOMA213; 22; 10. DOMA214; 22; 10. DOMA215; 22; 10. DOMA216; 22; 10. DOMA217; 22; 10. DOMA218; 22; 10. DOMA219; 22; 10. DOMA220; 22; 10. DOMA221; 22; 10. DOMA222; 22; 10. DOMA223; 22; 10. DOMA224; 22; 10. DOMA225; 22; 10. DOMA226; 22; 10. DOMA227; 22; 10. DOMA228; 22; 10. DOMA229; 22; 10. DOMA230; 22; 10. DOMA231; 22; 10. DOMA232; 22; 10. DOMA233; 22; 10. DOMA234; 22; 10. DOMA235; 22; 10. DOMA236; 22; 10. DOMA237; 22; 10. DOMA238; 22; 10. DOMA239; 22; 10. DOMA240; 22; 10. DOMA241; 22; 10. DOMA242; 22; 10. DOMA243; 22; 10. DOMA244; 22; 10. DOMA245; 22; 10. DOMA246; 22; 10. DOMA247; 22; 10. DOMA248; 22; 10. DOMA249; 22; 10. DOMA250; 22; 10. DOMA251; 22; 10. DOMA252; 22; 10. DOMA253; 22; 10. DOMA254; 22; 10. DOMA255; 22; 10. DOMA256; 22; 10. DOMA257; 22; 10. DOMA258; 22; 10. DOMA259; 22; 10. DOMA260; 22; 10. DOMA261; 22; 10. DOMA262; 22; 10. DOMA263; 22; 10. DOMA264; 22; 10. DOMA265; 22; 10. DOMA266; 22; 10. DOMA267; 22; 10. DOMA268; 22; 10. DOMA269; 22; 10. DOMA270; 22; 10. DOMA271; 22; 10. DOMA272; 22; 10. DOMA273; 22; 10. DOMA274; 22; 10. DOMA275; 22; 10. DOMA276; 22; 10. DOMA277; 22; 10. DOMA278; 22; 10. DOMA279; 22; 10. DOMA280; 22; 10. DOMA281; 22; 10. DOMA282; 22; 10. DOMA283; 22; 10. DOMA284; 22; 10. DOMA285; 22; 10. DOMA286; 22; 10. DOMA287; 22; 10. DOMA288; 22; 10. DOMA289; 22; 10. DOMA290; 22; 10. DOMA291; 22; 10. DOMA292; 22; 10. DOMA293; 22; 10. DOMA294; 22; 10. DOMA295; 22; 10. DOMA296; 22; 10. DOMA297; 22; 10. DOMA298; 22; 10. DOMA299; 22; 10. DOMA300; 22; 10. DOMA301; 22; 10. DOMA302; 22; 10. DOMA303; 22; 10. DOMA304; 22; 10. DOMA305; 22; 10. DOMA306; 22; 10. DOMA307; 22; 10. DOMA308; 22; 10. DOMA309; 22; 10. DOMA310; 22; 10. DOMA311; 22; 10. DOMA312; 22; 10. DOMA313; 22; 10. DOMA314; 22; 10. DOMA315; 22; 10. DOMA316; 22; 10. DOMA317; 22; 10. DOMA318; 22; 10. DOMA319; 22; 10. DOMA320; 22; 10. DOMA321; 22; 10. DOMA322; 22; 10. DOMA323; 22; 10. DOMA324; 22; 10. DOMA325; 22; 10. DOMA326; 22; 10. DOMA327; 22; 10. DOMA328; 22; 10. DOMA329; 22; 10. DOMA330; 22; 10. DOMA331; 22; 10. DOMA332; 22; 10. DOMA333; 22; 10. DOMA334; 22; 10. DOMA335; 22; 10. DOMA336; 22; 10. DOMA337; 22; 10. DOMA338; 22; 10. DOMA339; 22; 10. DOMA340; 22; 10. DOMA341; 22; 10. DOMA342; 22; 10. DOMA343; 22; 10. DOMA344; 22; 10. DOMA345; 22; 10. DOMA346; 22; 10. DOMA347; 22; 10. DOMA348; 22; 10. DOMA349; 22; 10. DOMA350; 22; 10. DOMA351; 22; 10. DOMA352; 22; 10. DOMA353; 22; 10. DOMA354; 22; 10. DOMA355; 22; 10. DOMA356; 22; 10. DOMA357; 22; 10. DOMA358; 22; 10. DOMA359; 22; 10. DOMA360; 22; 10. DOMA361; 22; 10. DOMA362; 22; 10. DOMA363; 22; 10. DOMA364; 22; 10. DOMA365; 22; 10. DOMA366; 22; 10. DOMA367; 22; 10. DOMA368; 22; 10. DOMA369; 22; 10. DOMA370; 22; 10. DOMA371; 22; 10. DOMA372; 22; 10. DOMA373; 22; 10. DOMA374; 22; 10. DOMA375; 22; 10. DOMA376; 22; 10. DOMA377; 22; 10. DOMA378; 22; 10. DOMA379; 22; 10. DOMA380; 22; 10. DOMA381; 22; 10. DOMA382; 22; 10. DOMA383; 22; 10. DOMA384; 22; 10. DOMA385; 22; 10. DOMA386; 22; 10. DOMA387; 22; 10. DOMA388; 22; 10. DOMA389; 22; 10. DOMA390; 22; 10. DOMA391; 22; 10. DOMA392; 22; 10. DOMA393; 22; 10. DOMA394; 22; 10. DOMA395; 22; 10. DOMA396; 22; 10. DOMA397; 22; 10. DOMA398; 22; 10. DOMA399; 22; 10. DOMA400; 22; 10. DOMA401; 22; 10. DOMA402; 22; 10. DOMA403; 22; 10. DOMA404; 22; 10. DOMA405; 22; 10. DOMA406; 22; 10. DOMA407; 22; 10. DOMA408; 22; 10. DOMA409; 22; 10. DOMA410; 22; 10. DOMA411; 22; 10. DOMA412; 22; 10. DOMA413; 22; 10. DOMA414; 22; 10. DOMA415; 22; 10. DOMA416; 22; 10. DOMA417; 22; 10. DOMA418; 22; 10. DOMA419; 22; 10. DOMA420; 22; 10. DOMA421; 22; 10. DOMA422; 22; 10. DOMA423; 22; 10. DOMA424; 22; 10. DOMA425; 22; 10. DOMA426; 22; 10. DOMA427; 22; 10. DOMA428; 22; 10. DOMA429; 22; 10. DOMA430; 22; 10. DOMA431; 22; 10. DOMA432; 22; 10. DOMA433; 22; 10. DOMA434; 22; 10. DOMA435; 22; 10. DOMA436; 22; 10. DOMA437; 22; 10. DOMA438; 22; 10. DOMA439; 22; 10. DOMA440; 22; 10. DOMA441; 22; 10. DOMA442; 22; 10. DOMA443; 22; 10. DOMA444; 22; 10. DOMA445; 22; 10. DOMA446; 22; 10. DOMA447; 22; 10. DOMA448; 22; 10. DOMA449; 22; 10. DOMA450; 22; 10. DOMA451; 22; 10. DOMA452; 22; 10. DOMA453; 22; 10. DOMA454; 22; 10. DOMA455; 22; 10. DOMA456; 22; 10. DOMA457; 22; 10. DOMA458; 22; 10. DOMA459; 22; 10. DOMA460; 22; 10. DOMA461; 22; 10. DOMA462; 22; 10. DOMA463; 22; 10. DOMA464; 22; 10. DOMA465; 22; 10. DOMA466; 22; 10. DOMA467; 22; 10. DOMA468; 22; 10. DOMA469; 22; 10. DOMA470; 22; 10. DOMA471; 22; 10. DOMA472; 22; 10. DOMA473; 22; 10. DOMA474; 22; 10. DOMA475; 22; 10. DOMA476; 22; 10. DOMA477; 22; 10. DOMA478; 22; 10. DOMA479; 22; 10. DOMA480; 22; 10. DOMA481; 22; 10. DOMA482; 22; 10. DOMA483; 22; 10. DOMA484; 22; 10. DOMA485; 22; 10. DOMA486; 22; 10. DOMA487; 22; 10. DOMA488; 22; 10. DOMA489; 22; 10. DOMA490; 22; 10. DOMA491; 22; 10. DOMA492; 22; 10. DOMA493; 22; 10. DOMA494; 22; 10. DOMA495; 22; 10. DOMA496; 22; 10. DOMA497; 22; 10. DOMA498; 22; 10. DOMA499; 22; 10. DOMA500; 22; 10. DOMA501; 22; 10. DOMA502; 22; 10. DOMA503; 22; 10. DOMA504; 22; 10. DOMA505; 22; 10. DOMA506; 22; 10. DOMA507; 22; 10. DOMA508; 22; 10. DOMA509; 22; 10. DOMA510; 22; 10. DOMA511; 22; 10. DOMA512; 22; 10. DOMA513; 22; 10. DOMA514; 22; 10. DOMA515; 22; 10. DOMA516; 22; 10. DOMA517; 22; 10. DOMA518; 22; 10. DOMA519; 22; 10. DOMA520; 22; 10. DOMA521; 22; 10. DOMA522; 22; 10. DOMA523; 22; 10. DOMA524; 22; 10. DOMA525; 22; 10. DOMA526; 22; 10. DOMA527; 22; 10. DOMA528; 22; 10. DOMA529; 22; 10. DOMA530; 22; 10. DOMA531; 22; 10. DOMA532; 22; 10. DOMA533; 22; 10. DOMA534; 22; 10. DOMA535; 22; 10. DOMA536; 22; 10. DOMA537; 22; 10. DOMA538; 22; 10. DOMA539; 22; 10. DOMA540; 22; 10. DOMA541; 22; 10. DOMA542; 22; 10. DOMA543; 22; 10. DOMA544; 22; 10. DOMA545; 22; 10. DOMA546; 22; 10. DOMA547; 22; 10. DOMA548; 22; 10. DOMA549; 22; 10. DOMA550; 22; 10. DOMA551; 22; 10. DOMA552; 22; 10. DOMA553; 22; 10. DOMA554; 22; 10. DOMA555; 22; 10. DOMA556; 22; 10. DOMA557; 22; 10. DOMA558; 22; 10. DOMA559; 22; 10. DOMA560; 22; 10. DOMA561; 22; 10. DOMA562; 22; 10. DOMA563; 22; 10. DOMA564; 22; 10. DOMA565; 22; 10. DOMA566; 22; 10. DOMA567; 22; 10. DOMA568; 22; 10. DOMA569; 22; 10. DOMA570; 22; 10. DOMA571; 22; 10. DOMA572; 22; 10. DOMA573; 22; 10. DOMA574; 22; 10. DOMA575; 22; 10. DOMA576; 22; 10. DOMA577; 22; 10. DOMA578; 22; 10. DOMA579; 22; 10. DOMA580; 22; 10. DOMA581; 22; 10. DOMA582; 22; 10. DOMA583; 22; 10. DOMA584; 22; 10. DOMA585; 22; 10. DOMA586; 22; 10. DOMA587; 22; 10. DOMA588; 22; 10. DOMA589; 22; 10. DOMA590; 22; 10. DOMA591; 22; 10. DOMA592; 22; 10. DOMA593; 22; 10. DOMA594; 22; 10. DOMA595; 22; 10. DOMA596; 22; 10. DOMA597; 22; 10. DOMA598; 22; 10. DOMA599; 22; 10. DOMA600; 22; 10. DOMA601; 22; 10. DOMA602; 22; 10. DOMA603; 22; 10. DOMA604; 22; 10. DOMA605; 22; 10. DOMA606; 22; 10. DOMA607; 22; 10. DOMA608; 22; 10. DOMA609; 22; 10. DOMA610; 22; 10. DOMA611; 22; 10. DOMA612; 22; 10. DOMA613; 22; 10. DOMA614; 22; </p>
--	---

SMART; SM00355; 2nf_C2H2; 5;	CC	ZINC-FINGER IN-BETWEEN.	
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.	CC	-!- SIMILARITY: STRONG, TO HIVEP2.	
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
KW Developmental protein; Segmentation polarity protein; Zinc-finger;	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
KW Metal-binding; DNA-binding; Repeat; Nuclear protein.	CC	the European Bioinformatics Institute. There are no restrictions on its	
DOMAIN 451-603 ZINC FINGERS.	CC	use by non-profit institutions as long as its content is in no way	
FT ZN_FING 451 476 C2H2-TYPE.	CC	modified and this statement is not removed. Usage by and for commercial	
FT ZN_FING 484 511 C2H2-TYPE.	CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
FT ZN_FING 517 541 C2H2-TYPE.	CC		
FT ZN_FING 547 572 C2H2-TYPE.	CC		
FT ZN_FING 578 603 C2H2-TYPE.	CC		
SQ SEQUENCE 1377 AA; 150881 MW; 85189314D5E7B3B7 CRC64;	CC		
Query Match 91.7%; Score 22; DB 1; Length 1377;	CC		
Best Local Similarity 66.7%; Pred. No. 3e+02; Mismatches 0; Gaps 0;	CC		
Matches 4; Conservative 0; Indels 0; Gaps 0;	CC		
QY 1 gtxxps 6	DR	EMBL; X51435; CAA35798.1; -	
Db 655 GtSSPS 660	DR	PIR; A34203; A34203.	
RESULT 27	DR	PDB; 3ZNF; 15-JAN-92.	
ZEP1_HUMAN STANDARD: PRT; 2717 AA.	DR	PDB; 4ZNF; 15-JAN-92.	
ID P15822;	DR	PDB; 1BBO; 31-OCT-93.	
AC P15822;	DR	TRANSFAC; T000497; -.	
DT 01-APR-1990 (Rel. 14, Created)	DR	MIM; 194540; -.	
DT 01-APR-1990 (Rel. 14, Last sequence update)	DR	InterPro; IPR00822; znf_C2H2.	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	DR	Pfam; PF00096; zf_C2H2; 5.	
DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-binding protein 1) (HIV-EP1) (Major histocompatibility complex binding protein 1) (MBP-1) (Positive regulatory domain III binding factor 1)	DR	PRINTS; PR00048; ZINC_FINGER_C2H2_4.	
DE (PRDI-B1).	DR	SMART; SM00355; Znf_C2H2_4.	
DE HIVEP1 OR ZNF40.	DR	PROSITE; PS50157; ZINC_FINGER_C2H2_1; 4.	
OS Homo sapiens (Human).	DR	PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	DR	KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	DR	KW Nuclear protein; Repeat; 3D-structure.	
NCBI_TAXID=9606; [1]	DR	DOMAIN 406 456 ZINC FINGERS.	
SEQUENCE FROM N.A. MEDLINE:9016514; PubMed=2106471;	FT	ZN_FING 406 428 C2H2-FINGERS.	
RA Fan C.M., Maniatis T.; containing two widely separated zinc finger motifs that recognize the same DNA sequence.";	FT	ZN_FING 434 456 C2H2-TYPE.	
RT [2]	FT	ZN_FING 803 806 POLY-SER.	
RT [2]	FT	ZN_FING 958 981 C2H2-TYPE (POTENTIAL).	
RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K., Gronenborn A.M.	FT	DOMAIN 2087 2139 ZINC FINGERS.	
RA "High-resolution three-dimensional structure of a single zinc finger from a human enhancer binding protein in solution.";	FT	ZN_FING 2115 2139 C2H2-TYPE.	
RA Biochemistry 29:9324-9334(1990).	FT	STRAND 2088 2088 C2H2-TYPE.	
RA [3]	FT	TURN 2090 2095 2095 2095	
RP STRUCTURE BY NMR OF 2113-2142. MEDLINE:91064333; PubMed=2248949;	FT	STRAND 2095 2108 2108 2108	
RX Omichinski J.G., Clore G.M., Appella E., Sakaguchi K., Gronenborn A.M.	FT	HELIX 2109 2109 2109 2109	
RA "High-resolution solution structure of the double Cys2His2 zinc finger from the human enhancer binding protein MBP-1.";	FT	TURN 2115 2116 2116 2116	
RA Biochemistry 31:3917-3917(1992).	FT	STRAND 2123 2124 2124 2124	
RT -!- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE 5'-GGGACTTCC-3', WHICH IS FOUND IN THE ENHANCER ELEMENTS OF NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1. IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT IN T-CELL ACTIVATION.	FT	HELIX 2127 2135 2135 2135	
RL RN	SEQUENCE 2717 AA; 297217 MW; D45D3CA951FBA561 CRC64;	FT	STRAND 2717 AA; 297217 MW; D45D3CA951FBA561 CRC64;
RT [1]	RESULT 28	Query Match 91.7%; Score 22; DB 1; Length 2717;	
RT [2]	THY_G_BOVIN STANDARD; PRT; 2769 AA.	Best Local Similarity 66.7%; Pred. No. 5.8e+02;	
RT [2]	ID THY_G_BOVIN ID P01267; Q18976; Q28196;	Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
RT [2]	AC AC DT 21-JUL-1986 (Rel. 01, Created)		
RT [2]	DT 21-JUL-1986 (Rel. 01, Last sequence update)		
RT [2]	DT 16-Oct-2001 (Rel. 40, Last annotation update)		
RT [2]	DE Thryglobulin Precursor. TN. GN OS Bos taurus (Bovine)		
RT [2]	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Bovidae; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.		
RT [2]	OC NCBI_TAXID=9913; [1]		
RT [2]	RP SEQUENCE FROM N.A. MEDLINE=85296288; PubMed=3855243;		
RT [2]	RX [1]		
RT [2]	RA Merchen L., Simon M.-J., Swillens S., Massart G.		
RT [2]	RA "Primary structure of bovine thyroglobulin deduced from the sequence of a cDNA clone." [1]		

of its 8,431-base complementary DNA.";

[2] Nature 316:647-651(1985).

RL REPEAT 1489 1505 TYPE II.

RN DOMAIN 1513 1567 THYROGLOBULIN TYPE IB 2.

RP REPEAT 1605 1725 TYPE IIIA.

SEQUENCE OF 1-930 FROM N.A.

RX MEDLINE=97190432; PubMed=3032624;

RP MEDLINE=05127025; PubMed=3855750;

RA Mercken L., Simons M.-J., de Martynhoff G., Swillens S., Vassart G.;

RT "Presence of hormonogenic and repetitive domains in the first 930 amino acids of bovine thyroglobulin as deduced from the cDNA sequence.";

RT Eur. J. Biochem. 147:59-64(1985).

RL REPEAT 1726 1893 IODINATION (IN T4).

RN [3] REPEAT 1794 1996 IODINATION (IN T4).

RP SEQUENCE OF 1-22 FROM N.A.

RX MEDLINE=07190432; PubMed=3032624;

RA de Martynhoff G., Pohl V., Mercken L., van Ommen G.-J., Vassart G.;

RT "Structural organization of the bovine thyroglobulin gene and of its 5'-flanking region.";

RL Eur. J. Biochem. 164:591-599(1987).

RN REPEAT 1997 2130 IODINATION (IN T4).

RP SEQUENCE OF 1002-1209 FROM N.A.

RX MEDLINE=88062712; PubMed=3381978;

RA Parma J., Christophe D., Pohl V., Vassart G.;

RT "Structural organization of the 5' region of the thyroglobulin gene. Evidence for intron loss and 'exonization' during evolution.";

J. Mol. Biol. 196:769-779(1987).

RL REPEAT 2131 2188 IODINATION (IN T4).

CC -I- PRECURSOR OF THE IODINATED THYROID HORMONES THYROXINE (T4) AND TRIIODOTHYRONINE (T3).

CC -I- SUBUNIT: HOMODIMER.

CC -I- SUBCELLULAR LOCATION: Secreted.

CC -I- TISSUE SPECIFICITY: THYROID GLAND SPECIFIC.

CC -I- PIM: SULFATED (BY SIMILARITY).

CC -I- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.

CC -I- SIMILARITY: CONTAINS 11 THYROGLOBULIN TYPE-I DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; X02815; CAA26584.1; .

DR EMBL; X012155; CAA16090.1; .

DR EMBL; X05380; CAA28971.1; ALT_SEQ.

DR EMBL; X06071; CAA29457.1; .

DR EMBL; X06072; CAA19457.1; JOINED.

DR EMBL; X06073; CAA29457.1; JOINED.

DR EMBL; X06074; CAA29457.1; JOINED.

DR EMBL; X06075; CAA19457.1; JOINED.

PTR; A01533; U1BO.

HSSP; P21836; 1MAR.

DR InterPro; IPR002018; Carboxyesterase_B.

DR InterPro; IPR002016; Thyroglobulin_1.

DR Pfam; PF00135; Coesterase_2.

DR SMART; SM00211; TY_10.

DR PROSITE; PS00941; CARBOXYLESTERASE_B-1.

KW Glycoprotein; Repeat; Thyroid hormone; Iodination; Sulfation; Signal.

FT SIGNAL 1 19

FT CHAIN 20 2769 THYROGLOBULIN.

FT DOMAIN 31 92 THYROGLOBULIN TYPE IA 1.

FT DOMAIN 93 160 THYROGLOBULIN TYPE IA 2.

FT DOMAIN 161 297 THYROGLOBULIN TYPE IA 3.

FT DOMAIN 298 358 THYROGLOBULIN TYPE IA 4.

FT DOMAIN 604 657 THYROGLOBULIN TYPE IA 5.

FT DOMAIN 661 725 THYROGLOBULIN TYPE IA 6.

FT DOMAIN 726 921 THYROGLOBULIN TYPE IA 7.

FT DOMAIN 922 1073 THYROGLOBULIN TYPE IA 8.

FT DOMAIN 1074 1145 THYROGLOBULIN TYPE IB 1.

FT DOMAIN 1146 1210 THYROGLOBULIN TYPE IA 9.

FT REPEAT 1458 1471 TYPE II.

FT REPEAT 1472 1488 TYPE II.

FT MOD_RES 24 24 IODINATION (IN T4), SULFATION (POTENTIAL).

FT MOD_RES 24 24 IODINATION (IN T4).

FT MOD_RES 2574 2574 IODINATION (IN T4).

FT MOD_RES 2588 2588 IODINATION (IN T3).

FT MOD_RES 2767 2767 IODINATION (IN T4).

FT CARBOHYD 110 110 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 198 198 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 483 483 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 495 495 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 747 747 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 853 853 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 947 947 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 1140 1140 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 1365 1365 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 1776 1776 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 1870 1870 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 2014 2014 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 2123 2123 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 2251 2251 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CARBOHYD 2296 2296 N-LINKED (GLCNAC . .) (POTENTIAL).

FT CONFLICT 1206 1206 S -> R (IN REF. 4).

SQ 2/69 AA; 303218 MW; 1IC7F227E9101DE2A CRC64;

Query Match 91.7%; Score 22; DB 1; Length 2769;

Best Local Similarity 66.7%; Pred. No. 5.9e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

SQ SEQUENCE 2/69 2196 GNSSPS 2201

RESULT 29

LYS3_ECOLI

ID LYS3_ECOLI STANDARD

AC P05821;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Lysis protein for colicin E1 precursor.

GN LYS.

OS Escherichia coli.

OG Plasmid ColE1.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia

NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N. A.

RX MEDLINE=8604231; PubMed=3936034;

RX RT "Structural and functional organization of the colicin E1 operon.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:11276-11280(1994).

CC -1- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE AND PARTIAL CELL LYSIS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC DR EMBL; M12543; AAA23067.1; -.
 CC DR PIR; B24685; ZHECP1; -.
 CC DR InterPro; IPR003059; Lysis_col.
 CC PFam; PF02402; Lysis_col; 1.
 CC PRINTS; PRO01297; LYSSCOLICIN.

CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Plasmid; Outer membrane; Lipoprotein; Signal.

CC KW SIGNAL 1 17 POTENTIAL.
 CC FT CHAIN 18 45 LYSS PROTEIN FOR COLICIN E1.
 CC FT LIPID 18 18 N-ACYL GLYLCERIDE (POTENTIAL).

CC SEQUENCE 45 AA; 4829 NW; A34D3B4FBC12A13E CRC64;
 CC OS *Escherichia coli*.
 CC OS Plasmid ColeE2-P9, and Plasmid ColeE3-CA38
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC OC Escherichia.
 CC OX NCBI_TaxID:562;

Query Match 87.5%; Score 21; DB 1; Length 45;
 Best Local Similarity 66.7%; Pred. No. 23; 2; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CC Qy 1 gtxxps 6
 CC Db 29 GTIAPS 34

RESULT 30
 ID LYS3_SHISO STANDARD; PRT; 45 AA.
 AC P21185;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysis protein for colicin E1* precursor.
 GN KIL.

CC OS *Shigella sonnei*.
 CC OS Plasmid pK7-1.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC OC Shigella.
 CC OX NCBI_TaxID:624;

CC RN SEQUENCE FROM N.A.
 CC RA Higashi M., Hata M., Hase T., Yamaguchi K., Masamune Y.;
 CC RT "The nucleotide sequence of cca and the region of origin of plasmid
 CC RT pK7-1.";
 CC RT J. Gen. Appl. Microbiol. 32:433-442(1986).
 CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
 CC AND PARTIAL CELL LYSIS.

CC -!- SIMILARITY: 96% IDENTITY TO E. COLI LYSIS PROTEIN PRECURSOR.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC DR EMBL; M37218; AAA98158.1; -.
 CC DR PIR; S10921; S10921.

CC DR InterPro; IPR001059; Lysis_col.
 CC DR PFam; PF02402; Lysis_col; 1.
 CC DR PRINTS; PRO01297; LYSSCOLICIN.

CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 CC KW Plasmid; Outer membrane; Lipoprotein; Signal.

CC FT SIGNAL 1 17 POTENTIAL.
 CC FT CHAIN 18 45 LYSS PROTEIN FOR COLICIN E1*.
 CC FT LIPID 18 18 N-ACYL GLYLCERIDE (POTENTIAL).

CC SEQUENCE 45 AA; 4831 NW; A34D32FBC12B99D CRC64; -.

Query Match 87.5%; Score 21; DB 1; Length 45;
 Best Local Similarity 66.7%; Pred. No. 23; 0; Mismatches 2; Indels 0; Gaps 0;

CC Qy 1 gtxxps 6
 CC Db 29 GTVAPS 34

RESULT 31
 ID LYS2_ECOLI STANDARD; PRT; 47 AA.

AC P06963;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Lysis protein for colicins E2 and E3 precursor.
 GN HIC OR CELB.
 OS *Escherichia coli*.
 OS Plasmid ColeE2-P9, and Plasmid ColeE3-CA38

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID:562;

RN [1]
 RP SEQUENCE FROM N.A.

RC PLASMID=COLE2-P9;
 RX MEDLINE=86195936; PubMed=3516985;

RA Toba M., Masaki H., Ohta T.;
 RA Cole S.T., Saint-Joannis B., Pugsley A.P.;
 RT "Primary structures of the ColeE2-P9 and ColeE3-CA38 lysis genes.";
 RT "Molecular characterisation of the colicin E2 operon and
 RT identification of its products.";
 RL J. Biochem. 99:591-596(1986).
 RN [12]
 RP SEQUENCE FROM N.A.

RC PLASMID=COLE2-P9;
 RX MEDLINE=85339907; PubMed=3892228;

RA Cole S.T., Saint-Joannis B., Pugsley A.P.;
 RT "Primary structures of the ColeE2-P9 and ColeE3-CA38 lysis genes.";

RL J. Biochem. 99:591-596(1986).
 RN [13]
 RP SEQUENCE FROM N.A.

RC PLASMID=COLE3-CA38;
 RX MEDLINE=85028427; PubMed=6092219;

RA Watson R.J., Lau P.C.K., Vernet T., Visentini L.P.;
 RT "Characterization and nucleotide sequence of a colicin-like gene
 RT in the hic region of Plasmid ColeE3-CA38.";
 RL Gene 29:175-184 (1984).
 RN [4]
 RP ERRATUM.
 RC PLASMID=COLE3-CA38;
 RA Watson R.J., Lau P.C.K., Vernet T., Visentini L.P.;
 RL Gene 42:351-353(1986).
 RN [5]
 RP SEQUENCE OF 1-38 FROM N.A.

RC PLASMID=COLE3-CA38;
 RX MEDLINE=8510906; PubMed=3889348;

RA Masaki H., Ohta T.;
 RT "Colicin E3 and its immunity genes.";
 RL J. Mol. Biol. 182:217-227(1985).
 CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
 CC AND PARTIAL CELL LYSIS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC DR PIR; S10921; S10921.

CC DR InterPro; IPR001059; Lysis_col.

CC DR PFam; PF02402; Lysis_col; 1.

CC DR PRINTS; PRO01297; LYSSCOLICIN.

CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.

CC KW Plasmid; Outer membrane; Lipoprotein; Signal.

CC FT SIGNAL 1 17 POTENTIAL.

CC FT CHAIN 18 45 LYSS PROTEIN FOR COLICIN E1*.

CC FT LIPID 18 18 N-ACYL GLYLCERIDE (POTENTIAL).

CC SEQUENCE 45 AA; 4831 NW; A34D32FBC12B99D CRC64; -.

CC DR X02337; CAA6243.1; -.

CC DR D00020; BAA00014.1; -.

DR EMBL; D00021; BAA00015.1; -.
 DR EMBL; J01574; AAA88419.1; -.
 DR EMBL; X03631; CAA27281.1; -.
 DR EMBL; X01632; CAA27282.1; -.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Plasmid; Outer membrane; Lipoprotein; Signal.
 DR EMBL; M29885; AAA23070.1; -.
 DR PIR; A22383; BVECH8.
 DR PIR; JS0004; JS0004.
 DR PIR; S09329; S09529.
 DR InterPro; IPR00059; Lysis_col.
 DR Pfam; PF02402; Lysis_col; 1.
 DR PRINTS; PR01297; LYSISOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Plasmid; Outer membrane; Lipoprotein; Signal.
 KW POTENTIAL.
 FT SIGNAL 1 19
 FT CHAIN 20 47
 FT LIPID 20 20
 SQ SEQUENCE 47 AA; 4926 MW; 69133B25CA15A4B8 CRC64;
 DR InterPro; IPR00059; Lysis_col.
 DR Pfam; PF02402; Lysis_col; 1.
 DR PRINTS; PR01297; LYSISOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Plasmid; Outer membrane; Lipoprotein; Signal.
 KW POTENTIAL.
 FT SIGNAL 1 19
 FT CHAIN 20 47
 FT LIPID 20 20
 SQ SEQUENCE 47 AA; 4926 MW; 69133B25CA15A4B8 CRC64;
 DR InterPro; IPR00059; Lysis_col.
 DR Pfam; PF02402; Lysis_col; 1.
 DR PRINTS; PR01297; LYSISOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Plasmid; Outer membrane; Lipoprotein; Signal.
 DR EMBL; M29885; AAA23070.1; -.
 DR PIR; J00330; J00330.
 DR PIR; CA5799; C45799.
 DR InterPro; IPR003039; Lysis_col.

DR Pfam; PF02402; Lysis_col; 1.
 DR PRINTS; PR01297; LYSISOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Plasmid; Outer membrane; Lipoprotein; Signal.
 KW POTENTIAL.
 FT SIGNAL 1 19
 FT CHAIN 20 47
 FT LIPID 20 20
 SQ SEQUENCE 47 AA; 4926 MW; 69133B25CA15A4B8 CRC64;
 DR InterPro; IPR00059; Lysis_col.
 DR Pfam; PF02402; Lysis_col; 1.
 DR PRINTS; PR01297; LYSISOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Plasmid; Outer membrane; Lipoprotein; Signal.
 DR EMBL; M29885; AAA23070.1; -.
 DR PIR; J00330; J00330.
 DR PIR; CA5799; C45799.
 DR InterPro; IPR003039; Lysis_col.

Query Match 87.5%; Score 21; DB 1; Length 47;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 31 GTVSPS 36

RESULT 33
 LYS5_ECOLI STANDARD; PRT; 47 AA.
 ID LYS5_ECOLI
 AC P13345;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Lysis protein for colicin E6 precursor.
 DE Lysin.
 OC Escherichia coli.
 OS Escherichia coli.
 OC Plasmid Cole6-CT14.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90078082; PubMed=2687234;
 RA Akutsu A., Masaki H., Ohta T.;
 RA Lau P.C.K., Condie J.A.;
 RA "Nucleotide sequences from the colicin E5, E6 and E9 operons; presence of a degenerate transposon-like structure in the ColE9-J plasmid.";
 RT Mol. Gen. Genet. 217:269-277(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90078082; PubMed=2687234;
 RA Akutsu A., Masaki H., Ohta T.;
 RA "Molecular structure and immunity specificity of colicin E6, an evolutionary intermediate between E-group colicins and Cloacin DF13.";
 RT J. Bacteriol. 171:6430-6436(1989).
 RL CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE AND PARTIAL CELL LYSIS.
 CC
 RP SEQUENCE FROM N.A.
 RX MEDLINE-902218006; PubMed=25661131;
 RA Curtis M.D., James R., Coddington A.;
 RA "An evolutionary relationship between the ColE5-099 and the Cole9-J plasmids revealed by nucleotide sequencing.";
 RA J. Gen. Microbiol. 135:2783-2788(1989).
 RL Mol. Gen. Genet. 217:269-277(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89364708; PubMed=2549375;
 RA Lau P.C.K., Condie J.A.;
 RA "Nucleotide sequences from the colicin E5, E6 and E9 operons; presence of a degenerate transposon-like structure in the Cole9-J plasmid.";
 RA J. Gen. Microbiol. 135:2783-2788(1989).
 RL CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE AND PARTIAL CELL LYSIS.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X15856; CAA33858.1; -.
 CC DR EMBL; M31808; AAA23083.1; -.
 CC DR PIR; J00328; JQ0328.
 CC DR InterPro; IPR003059; Lysis_col.
 CC DR Pfam; PF02402; Lysis_col; 1.
 CC DR PRINTS; PR01297; LYSISOLICIN.
 CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Plasmid; Outer membrane; Lipoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 47
 FT LIPID 20 20
 SQ SEQUENCE 47 AA; 4902 MW; 691E124F6A057ED8 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 47;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 31 GtVSPS 36

RESULT 34
 LYS7_ECOLI STANDARD; PRT; 47 AA.
 ID LYS7_ECOLI
 AC Q03799;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysis protein for colicin E7 precursor.
 GN OS
 EScherichia coli.
 Plasmid ColeE7.
 OG Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia coli.
 OC Plasmid ColeE7.
 OC Escherichia coli.
 OX NCBI_TaxID=562;
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88121677; PubMed=3323826;
 RA Uchiumura T.; Lau P.C.K.;
 RT "Nucleotide sequences from the colicin E8 operon: homology with
 Plasmid ColeE2-P9.";
 RL Mol. Gen. Genet. 209:489-493 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88257046; PubMed=3290201;
 RA Toba M.; Masaki H.; Ohta T.;
 RT "Colicin E8, a DNase which indicates an evolutionary relationship
 between colicins E2 and E3.";
 RL J. Bacteriol. 170:1237-1242 (1988).
 CC -!- FUNCTION: LYSIS PROTEINS ARE REQUIRED FOR BOTH COLICIN RELEASE
 AND PARTIAL CELL LYSIS.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC
 DR M21404; AAA23075.1;
 DR EMBL; X06119; CAA23493.1;
 DR PIR; C2818; ZHECE8.
 DR PIR; S01082; S01082.
 DR InterPro; IPR003059; Lysis_col.
 DR Pfam; PF00202; Lysis_col; 1.
 DR PRINTS; PRO1297; LYSISCOLICIN.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Plasmid; Outer membrane; Lipoprotein; Signal.
 KW FT SIGNAL 1 19
 FT CHAIN 20 47 LYSIS PROTEIN FOR COLICIN E8.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).
 SQ SEQUENCE 47 AA; 4844 MW; 691E149ABA056B3B CRC64;

Query Match 87.5%; Score 21; DB 1; Length 47;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 31 GtVSPS 36

RESULT 36
 LYS9_ECOLI STANDARD; PRT; 47 AA.
 ID LYS9_ECOLI
 AC P15176;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-Oct-2001 (Rel. 40, Last annotation update)
 DE Lysis protein for colicin E9 precursor.
 GN LYS
 EScherichia coli.
 OG Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia coli.
 OC Plasmid ColeE9-J.
 OC Escherichia coli.
 OX NCBI_TaxID=562;

DR PRINTS; PRO1297; LYSISCOLICIN.
 KW PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Plasmid; Outer membrane; Lipoprotein; Signal.
 FT SIGNAL 1 17 LYSIS PROTEIN FOR COLICIN N.
 FT CHAIN 1 52 LYSIS PROTEIN FOR COLICIN N.
 FT LIPID 1 18 N-ACYL DIGLYCERIDE.
 SQ SEQUENCE 52 AA; 5633 MW; F6F1BF59BDAF81B6 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 52;
 Best Local Similarity 66.7%; Pred. No. 26;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0; Gaps 0;

Qy 1 gtxxps 6
 Db 29 GTVAPS 34

RESULT 39
 ID PAHO_CHICK STANDARD; PRT; 80 AA.
 AC P01306;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pancreatic hormone precursor (Pancreatic polypeptide) (PP).
 OS Gallus gallus (Chicken), and
 OS Meleagris gallopavo (Common turkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC NCBII_TAXID=9031, 9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Chicken; TISSUE=Liver;
 RX MEDLINE=93366173; PubMed=835965;
 RA Nata K., Sugimoto T., Konri K., Hidaka H., Hattori E., Yamamoto H.,
 RA Yonekura H., Okamoto H.;
 RT "Structure determination and evolution of the chicken cDNA and gene
 RT encoding prepropancreatic polypeptide.";
 RN Gene 130:183-189(1993).
 RN [2]
 RP SEQUENCE OF 26-61.
 RC SPECIES=Chicken;
 RX MEDLINE=76069270; PubMed=1194289;
 RA Kimmel J.R., Hayden L.J., Pollock H.G.;
 RT "Isolation and characterization of a new pancreatic polypeptide
 RT hormone.";
 RL J. Biol. Chem. 250:9369-9376(1975).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROEMS).
 RC SPECIES=M_gallopavo;
 RX MEDLINE=84179397; PubMed=6673760;
 RA Glover I., Maneef I., Pitts J., Woods S., Moss D., Tickle I.,
 RA Blundell T.L.;
 RT "Conformational flexibility in a small globular hormone: X-ray
 RT analysis of avian pancreatic polypeptide at 0.98-A resolution.";
 RL Biopolymers 22:293-304(1993).
 CC 1- FUNCTION: PANCREATIC HORMONE IS SYNTHESIZED IN PANCREATIC ISLETS
 CC OF LANGERHANS AND ACTS AS A REGULATOR OF PANCREATIC AND
 CC GASTROINTESTINAL FUNCTIONS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE NPY / PYY FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D13760; BAA02906.1; -.
 DR PIR; A01575; PCCH.
 DR PDB; 1PPT; 15-OCT-91.
 DR InterPro; IPR001955; Pancreatic hormone.
 DR Pfam; PF00159; Pancreatic hormone.
 DR PRINTS; PRO00278; PANHORMONE.
 DR PRODOM; P0001267; Pancreatic hormone; 1.
 DR SMART; SM00309; PAH; 1.
 DR PROSITE; PS00265; PANCREATIC HORMONE; 1.
 DR FROSTIE; PS50276; PANCREATIC_HORMONE_2; 1.
 KW Hormone; Cleavage on pair of basic residues; Pancreas; Signal;
 KW Anidation; 3D-structure.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 61 PANCREATIC HORMONE.
 FT MOD_RES 61 61 AMIDATION (G-62 PROVIDE AMIDE GROUP).
 FT CONFLICT 47 48 ND -> DN (IN REF. 2).
 FT TURN 35 35
 FT HELIX 39 56
 FT TURN 57 58
 SQ SEQUENCE 80 AA; 8773 MW; 90B44E27389DB050 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 80;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 23 GTAGPS 28

RESULT 40
 YGV_BACSU STANDARD; PRT; 94 AA.
 ID YGV_BACSU STANDARD;
 AC P54499;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 10.5 kDa protein in SODA-CONGA intergenic region.
 GN YGV.
 OC Bacteria; Firmicutes; Clostridium group; Bacillus.
 OC Bacillus subtilis.
 OC Bacillus/Staphylococcus group; Bacillus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH64;
 RA Kobayashi Y., Mizuno M., Masuda S., Takenaru K., Hcosono S.,
 RA Sato T., Takeuchi M.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE UPF0045 FAMILY.
 CC -----
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D84433; BAA12525.1; -.
 DR PIR; Z99116; CAB1412.1; -.
 DR Sublist; BG11689; YggV.
 DR InterPro; IPR002767; DUF77.
 DR Pfam; PF0110; DUF77; 1.
 KW Hypothetical protein. Complete proteome.
 SQ SEQUENCE 94 AA; 10453 MW; 4A57612338B80999 CRC64;

Query Match 87.5%; Score 21; DB 1; Length 94;
 Best Local Similarity 66.7%; Pred. No. 46;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DR EMBL; D13761; BAA02907.1; -.

QY 1 gtxxps 6
 Db 12 GTEPPS 17

RESULT 4.1

TRH4_ECOLI STANDARD; PRT; 118 AA.

ID TRH4_ECOLI
 AC Q00190;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE TRAH protein.
 GN TRAH.
 OS Escherichia coli.
 OG Plasmid IncP-beta RP4.
 OC Bacteriophage; gamma subdivision; Enterobacteriaceae;
 OC Escherichia; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia;
 OX NCBI_TaxID=562;

RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5.
 RC STRAIN=HB1101;
 RX MEDLINE=92190548; PubMed=1665997;
 RA Ziegellin G.; Pansegrouw W.; Strack B.; Balzer D.; Kroeger M.;
 RA Kruit V.; Lankau E.;
 RA "Nucleotide sequence and organization of genes flanking the transfer origin of promiscuous plasmid RP4.";
 RL DNA Seq. 1:303-327 (1991).
 -: FUNCTION: THE INITIATION PROCESS OF TRANSFER DNA SYNTHESIS
 CC REQUIRES THE INTERACTION OF AT LEAST THREE PLASMID-SPECIFIC
 CC COMPONENTS (TRAH, I, AND J) AT THE TRANSFER ORIGIN RESULTING
 CC IN THE ASSEMBLY OF A SPECIALISED NUCLEOPROTEIN COMPLEX - THE
 CC RELAXOSOME.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

QY 1 gtxxps 6
 Db 37 GTLAPS 42

Query Match 87.5%; Score 21; DB 1; Length 122;
 Best Local Similarity 66.7%; Pred. No. 57;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 4.3

FMF7_ECOLI STANDARD; PRT; 170 AA.

ID FMF7_ECOLI
 AC P23394;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-JUL-1993 (Rel. 22, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE FMF7_fimbrial protein precursor.
 GN FMFA.
 OS Escherichia coli.
 OC Bacteriophage; gamma subdivision; Enterobacteriaceae;
 OC Escherichia; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC NCBI_TaxID=562;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=107/86;
 RX MEDLINE=9222617; PubMed=1148723;
 RA Imberechts H.A.; de Geve H.; Schlicker C.; Bouchet H.; Pohl P.;
 RA Charlier G.; Vandekerckhove J.; van Montagu M.;
 RA Lintermans P.;
 RT "Characterization of F107 fimbriae of Escherichia coli 107/86, which
 RT causes edema disease in pigs, and nucleotide sequence of the F107
 RT major fimbrial subunit gene, fimbA.";
 RL Infect. Immun. 60:1963-1971 (1992).
 -: FUNCTION: FIBRILLAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
 CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
 CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
 CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
 CC -1 - SUBCELLULAR LOCATION: Fimbria.

Query Match 87.5%; Score 21; DB 1; Length 118;
 Best Local Similarity 66.7%; Pred. No. 57;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 4.2

RK14_PORPU STANDARD; PRT; 122 AA.

ID RK14_PORPU
 AC P51304;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE Chloroplast 50S ribosomal protein L14.
 GN Porphyra purpurea.
 OS Chloroplast; Chlorophyll a/b binding protein.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
 OX NCBI_TaxID=2787;
 RN [1]
 RP SEQUENCE FROM N.A.

DR	EMBL; M61713; AAA23735.1;	;	;	;	;	;
DR	M61713; AAA23734.1;	;	;	;	;	;
DR	PIR; A43841; A43841;	;	;	;	;	;
DR	InterPro; IPR000559; Fimbrial.	;	;	;	;	;
DR	Pfam; PF00419; Fimbrial;	1.	;	;	;	;
KW	Fimbria; Signal.	;	;	;	;	;
FT	SIGNAL	1	21	;	;	;
FT	CHAIN	22	170	F107 FIMBRIAL PROTEIN.	;	;
FT	DISULFID	37	78	PROBABLE.	;	;
SQ	SEQUENCE	170 AA:	17329 MW:	AF40947CE387692F CRC64;	;	;
RESULT 4						
ON24_ARATH						
ID	ON24_ARATH	STANDARD;	PRT;	187 AA.	;	;
AC	PB2805;	;	;	;	;	;
DT	16-OCT-2001 (Rel. 40, Created)	;	;	;	;	;
DT	16-OCT-2001 (Rel. 40, Last sequence update)	;	;	;	;	;
DT	16-OCT-2001 (Rel. 40, Last annotation update)	;	;	;	;	;
DE	Mitochondrial import receptor subunit TOM20-4 (Translocase of outer membrane 20 kDa subunit 4).	;	;	;	;	;
GN	TOM20-4 OR AT5G40930 OR MMC1.2.	;	;	;	;	;
OS	Arabidopsis thaliana (Mouse-ear cress).	;	;	;	;	;
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	;	;	;	;	;
OC	Spermatophyta; Magnoliophyta; eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	;	;	;	;	;
OC	NCBI_TAXID=3702;	;	;	;	;	;
RN	[1]	;	;	;	;	;
SEQUENCE FROM N A.						
RC	SEQUIN-CV, COLUMBIA;	;	;	;	;	;
RC	MEDLINE:21016721; PubMed=11130714;	;	;	;	;	;
RA	Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E., Miyajima N., Sasaki T., Kimura T., Hosokawa T., Kawashima K., Kohara M., Matsunoto M., Matsuno A., Muraki A., Nakayama S., Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M., Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R., Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J., Stonenkering T., Pepin K., Spiech L., Sekhon M., Armstrong J., Becker M., Beltaire E., Cordin H., Cordes M., Courtney W., Dante M., Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Lettreille P., Leonardi S., Meyer R., Mulvaney E., Ozerksy P., Riley A., Strowmatt C., Wagner-McPherson C., Wollam A., Youkum M., Dediha N., Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J., Kirchoff K., Roth K., King L., Bahret A., Miller B., Marr M., Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I., Volckaert G., Wainburt R., Duesterhoeft A., Stillekema W., Pohl T., Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S., Langham S.-A., McCullagh B., Robben J., Grymonpre B., Zimmermann W., Ransperger U., Wedler H., Balke K., Wedler E., Peters S., RA van Staveren M., Dirks W., Moijl P., Klein Lankhorst R., Weltzienegger T., Botha G., Rose M., Hauf J., Berneis S., Hempel S., RA Feldpausch M., Lambertz S., Villarroel R., Gielens J., Ardiles W., Bents O., Lencke K., Kolosov G., Mayer K.F.X., Rudd S., Schoof H., RA Schueler C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.; RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis thaliana;" RT Nature 408:823-826(2000).	;	;	;	;	;
RN	[2]	;	;	;	;	;
RP	SEQUIN-CV, COLUMBIA;	;	;	;	;	;
RC	published=11161051;	;	;	;	;	;
RC	Wernau W., Niemeyer A., Jaensch L., Kruft V., Schmitz U.K., Wernau H.-P.	;	;	;	;	;

entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

CC DR AE000037; AAB96035.1;
CC KW HYPOTHETICAL protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 10 30 POTENTIAL.
CC FT TRANSMEM 136 156 POTENTIAL.
CC SQ SEQUENCE 193 AA; 21443 MW; BAE6806C6D80C1D CRC64;

Query Match 87.5%; Score 21; DB 1; Length 193;
Best Local Similarity 66.7%; Pred. No. 91; 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	9txxps	6
Db	31	GTVSPS	36

RESULT 46

COX3_MYCTU	STANDARD;	PRT;	203 AA.
ID COX3_MYCTU			
AC Q10385;			
DT 01-OCT-1996 (Rel. 34, Created)			
DT 01-OCT-1996 (Rel. 34, Last sequence update)			
DT 16-OCT-2001 (Rel. 40, Last annotation update)			
DE Probable cytochrome c oxidase polypeptide III (EC 1.9.3.1) (Cytochrome aa3 subunit 3)			
DE CTAE OR RV2193 OR MTcY190.04.			
GN Mycobacterium tuberculosis.			
OC			
OC Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Actinomycetaceae; Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OC NCBI_TAXID:1773;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=H37RV;			
RX MEDLINE=902595987; PubMed=9634230;			
RA Cole S.T., Broosch R., Parkhill J., Garnier T., Churcher C., Harris D.,			
RA Gordon S.V., Eigmiller K., Gas S., Barry C.E. III, Tekkia F.,			
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,			
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,			
RA Hornsby T., Jagger K., Krogh A., McLean J., Moule S., Murphy L.,			
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,			
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,			
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,			
RA "Deciphering the biology of Mycobacterium tuberculosis from the			
RT complete genome sequence."; RT			
RT Nature 393:537-544 (1998).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=CDC 1551 / Oshkosh;			
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,			
RA Peterson J., Derby R., Dodson R., Ginn M.L., Haft D., Hickey E.,			
RA Kolonay J.F., Nelson W.C., Umayam L., Brimblebaeva M.D., Salzberg S.L.,			
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,			
RA Bishai W.;			
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and			
RT laboratory strains."; RT			
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
CC -i- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome			
CC C + 2 H(2)O			
CC -i- SUBCELLULAR LOCATION: Integral membrane protein.			
CC -i- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).			
CC DR Z70283; QAA94262; 1;			

RP SEQUENCE FROM N.A.
 STRAIN=MC58 / SEROGROUP B;
 MEDLINE=20175755; PubMed=10710307;

RC Tertetlin H., Saunders N.J., Heidelberger J., Jeffries A.C., Nelson K.E.,
 Eisens J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 Nelson W.C., Gwin M.L., Debay R., Peterson J.D., Hickey E.K.,
 Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
 Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 Gill J., Scarlato V., Masiolini V., Pizza M., Grandi G., Sun L.,
 Smith H.O., Fraser C.M., Moxon E.R., Rappoli R., Venter J.C.,
 RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
 MC58.";
 RL Science 287:1809-1815(2000).
 CC -1- FUNCTION: CATALYZES THE CONDENSATION OF 1-DEOXY-D-XYULOSE-5-
 CC PHOSPHATE (DXP) AND 1-AMINO-3-OXO-4-(PHOSPHOHYDROXY)PROPAN-2-ONE
 CC TO FORM PYRIDOXINE 5'-PHOSPHATE (PNP) (BY SIMILARITY).
 CC -1- PATHWAY: DE NOVO SYNTHESIS OF PYRIDOXINE (VITAMIN B6) AND
 CC PYRIDOXAL PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE PDXJ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use, by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR A5002401; AAF40885.1; -.
 DR TIGR; NMB0448; -.
 KW Pyridoxine biosynthesis; Complete proteome.

SQ SEQUENCE 242 AA; 25565 MW; 5DA0476728AA1485 CRC64;
 Query Match 87.5%; Score 21; DB 1; Length 242;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 18 GTTTPS 23

Search completed: September 3, 2002, 09:29:46
 Job time: 228 sec

Result No.	Score	Query Match	Length	DB ID	Description
1	22	91.7	71	Q9IF05	09if05 human immunopneumovirus
2	22	91.7	75	Q41118	Q41118 human immunopneumovirus
3	22	91.7	80	Q90QH8	Q90QH8 human immunopneumovirus
4	22	91.7	89	Q9HWZ4	Q9HWZ4 pseudomonas aeruginosa
5	22	91.7	92	Q38158	Q38158 human immunopneumovirus
6	22	91.7	93	Q38175	Q38175 human immunopneumovirus
7	22	91.7	93	Q38177	Q38177 human immunopneumovirus
8	22	91.7	96	Q38114	Q38114 human immunopneumovirus
9	22	91.7	96	Q38161	Q38161 human immunopneumovirus
10	22	91.7	96	Q38178	Q38178 human immunopneumovirus
11	22	91.7	99	Q38030	Q38030 human immunopneumovirus
12	22	91.7	99	Q38362	Q38362 human immunopneumovirus
13	22	91.7	99	Q38371	Q38371 human immunopneumovirus
14	22	91.7	99	Q38372	Q38372 human immunopneumovirus
15	22	91.7	101	Q38113	Q38113 human immunopneumovirus
16	22	91.7	101	Q38128	Q38128 human immunopneumovirus

Run on:	September 3, 2002, 09:25:33 ; Search time 25.06 Seconds (without alignments) 41.419 Million cell updates/sec	17	22	91.7	101	15	Q38130	human immunopneumovirus
Title:	BASK-853-CLAIM5	18	22	91.7	101	15	Q38133	human immunopneumovirus
Perfect score:	24	19	22	91.7	101	15	Q38138	human immunopneumovirus
Sequence:	1 gtxxps 6	20	22	91.7	101	15	Q38141	human immunopneumovirus
Scoring table:	BLOSUM62	21	22	91.7	105	17	Q9YA39	aeropyrum pneumoniae
Gapp:	10.0 , Gapext 0.5	22	22	91.7	111	15	Q38165	human immunopneumovirus
Searched:	562222 seqs, 172994929 residues	23	22	91.7	111	15	Q38166	human immunopneumovirus
Total number of hits satisfying chosen parameters:	562222	24	22	91.7	111	15	Q38169	human immunopneumovirus
Minimum DB seq length: 0		25	22	91.7	111	15	Q38171	human immunopneumovirus
Maximum DB seq length: 2000000000		26	22	91.7	112	15	Q75698	human immunopneumovirus
Post-processing: Minimum Match 0%		27	22	91.7	135	11	Q64053	maize rayado
Maximum Match 100%		28	22	91.7	142	11	Q9D5C6	mus musculus
Listing first 50 summaries		29	22	91.7	156	4	Q9HB7	homo sapiens
Database :	SPTREMBL_19:**	30	22	91.7	157	4	Q9WG3	rat cytomegalovirus
1:	sp_archaea:**	31	22	91.7	175	5	Q9WS07	streptomyces
2:	sp_bacteria:**	32	22	91.7	177	12	Q41995	maize rayado
3:	sp_fungi:**	33	22	91.7	178	2	Q9Y82	streptomyces
4:	sp_human:**	34	22	91.7	182	5	Q9WS06	streptomyces
5:	sp_invertebrate:**	35	22	91.7	186	4	Q9BZ94	homo sapiens
6:	sp_mammal:**	36	22	91.7	196	10	Q9B664	arabidopsis thaliana
7:	sp_mhc:**	37	22	91.7	200	2	Q9FS13	agrobacterium
8:	sp_organelle:**	38	22	91.7	202	16	Q9KQ11	vibrio cholerae
9:	sp_phage:**	39	22	91.7	218	11	Q9JLH5	rat
10:	sp_plant:**	40	22	91.7	228	2	Q9RDK2	streptomyces
11:	sp_rabbit:**	41	22	91.7	233	2	Q9KXV0	streptomyces
12:	sp_virus:**	42	22	91.7	237	6	P79289	sus scrofa
13:	sp_vertbrate:**	43	22	91.7	247	2	Q9L203	streptomyces
14:	sp_unclassified:**	44	22	91.7	253	2	Q91126	streptomyces
15:	sp_rvirus:**	45	22	91.7	253	4	Q43563	homo sapiens
16:	sp_bacteria:**	46	22	91.7	254	16	Q9i170	pseudomonas
17:	sp_archaea:**	47	22	91.7	275	13	Q98TY7	brachydorina
		48	22	91.7	285	10	Q9m0A9	arabidopsis thaliana
		49	22	91.7	288	2	Q9RNC9	barotella
		50	22	91.7	288	12	Q9dWd6	rat cytomegalovirus

ALIGNMENTS

RESULT	1	PRELIMINARY;	ERT;	71 AA.
ID	Q9IF05			
AC	Q9IF05;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	DEG PROTEIN (FRAGMENT).			
GN	POL.			
OS	Human immunodeficiency virus type 1			
OC	Viruses; Retroviruses; Lentiviridae; Lentivirus.			
RN	NCBI_TaxID=11676;			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE=1077;			
RX	MEDLINE=21443958; PubMed=11559796;			
RA	Perrin S., Munoz M., Yerly S., Sanchez-Merino V., Lopez-Galindez C., Perrin L., Larder B., Chmarko D., Fokan S., Meylan P., Telenti A.;			
RT	"Resistance to Nucleoside Analogs Reverse Transcriptase Inhibitors Mediated by Human Immunodeficiency Virus Type 1 P6 Protein."			
RL	J. Virol. 75:9614-9633 (2001).			
DR	EMBL: AF282969; AAF87830.1; -			
FT	NON_TER	1	1	
SQ	SEQUENCE 71 AA; 7870 MW; 033C6007FBCF3 CRC64;			
	[1]			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	91.7	71	Q9IF05	09if05 human immunopneumovirus
2	22	91.7	75	Q41118	Q41118 human immunopneumovirus
3	22	91.7	80	Q90QH8	Q90QH8 human immunopneumovirus
4	22	91.7	89	Q9HWZ4	Q9HWZ4 pseudomonas aeruginosa
5	22	91.7	92	Q38158	Q38158 human immunopneumovirus
6	22	91.7	93	Q38175	Q38175 human immunopneumovirus
7	22	91.7	93	Q38177	Q38177 human immunopneumovirus
8	22	91.7	96	Q38114	Q38114 human immunopneumovirus
9	22	91.7	96	Q38161	Q38161 human immunopneumovirus
10	22	91.7	96	Q38178	Q38178 human immunopneumovirus
11	22	91.7	99	Q38030	Q38030 human immunopneumovirus
12	22	91.7	99	Q38362	Q38362 human immunopneumovirus
13	22	91.7	99	Q38371	Q38371 human immunopneumovirus
14	22	91.7	99	Q38372	Q38372 human immunopneumovirus
15	22	91.7	101	Q38113	Q38113 human immunopneumovirus
16	22	91.7	101	Q38128	Q38128 human immunopneumovirus

Query Match	Local Similarity	Score 22;	DB 15;	Length 71;
Best_Matches	4;	Conservative	66.7%;	Pred. No. 1e+02;
Mismatches	2;	Indels	0;	Gaps 0;
QY	1	gttxxps 6		
Db	39	GTATPs 44		

RESULT	2	PRELIMINARY;	PRT;	75 AA.	RESULT	4	PRELIMINARY;	PRT;	89 AA.	
ID	Q41118				ID	Q9HWZ4				
AC	Q41118;	PRELIMINARY;	PRT;	75 AA.	AC	Q9HWZ4;				
DT	01-NOV-1996	(TREMBLrel. 01, Created)			DT	01-MAR-2001	(TREMBLrel. 16, Created)			
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)			DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			DT	01-OCT-2001	(TREMBLrel. 18, Last annotation update)			
DE	CHALCONE SYNTHASE mRNA (FRAGMENT).				DE	HYPOTHETICAL PROTEIN PA4033.				
OS	Phaseolus vulgaris (Kidney bean) (French bean).				GN	PA4033.				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.				OS	Pseudomonas aeruginosa.				
OC	NCBL-TaxID=3885;				OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;				
OX					OC	Pseudomonas				
RN					OX	NCBL-TaxID=387;				
RP					RN					
RA	SEQUENCE FROM N.A.				RP	SEQUENCE FROM N.A.				
RA	Ryder T. B., Cramer C. L., Bell J. N., Robbins M. P., Dixon R. A., Lamb C. J.;				RC	STRAIN=ATCC 15692 / PA01;				
RA	"Elicitor rapidly induces chalcone synthase mRNA in Phaseolus vulgaris cells at the onset of the phytoalexin defense response.";				RC	MEDLINE=20437337; PubMed=10984043;				
RT	Proc. Natl. Acad. Sci. U.S.A. 81:5724-5728(1984).				RA	Stover C. K., Pham X.-Q.-T., Erwin A. L., Mizoguchi S. D., Warrener P., Hickey M. J., Brinkman F. S. L., Hufnagle W. O., Kowalki D. J., Lagrou M., Garber R. L., Goltz L., Tolentino E., Westbroek-Wadman S., Yuan Y., Brody L. L., Coulter S. N., Folger K. R., Kas A., Larbig K., Lim R. M., Smith K. A., Spencer D. H., Wong G. K.-S., Wu Z., Paulsen I. T., Reizer J., Sauer M. H., Hancock R. E. W., Lory S., Olson M. V.;				
RL	EMBL; K02953; AAA33758.1;				RA	"Complete genome sequence of <i>Pseudomonas aeruginosa</i> PAO1, an opportunistic pathogen.";				
DR	InterPro; IPR001099; Chal_stil_synt.				RT	Opportunistic pathogen.";				
DR	InterPro; IPR001195; Chal_stil_synt_1.				RT	Opportunistic pathogen.";				
DR	Pfam; PF000199; Chal_stil_synt_1.				RL	Nature 406:959-964 (2000).				
FT	NON_TER 75 75				DR	AE004920; AAG07420.1; -.				
FT	NON_TER 75 75				KW	Hypothetical protein; Complete Protein.				
SEQUENCE	75 AA; 8800 MW;				SQ	SEQUENCE 99 AA; 9452 MW; BFD4EBF98897E4 CRC64;				
Qy	1 grtxxps 6				Query Match	91.7%; Score 22; DB 16; Length 89;				
Db	18 GTATPS 23				Best Local Similarity	91.7%; Score 22; DB 16; Length 89;				
Qy	1 grtxxps 6				Matches	4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
Db	18 GTATPS 23				Qy	• 1 grtxxps 6				
RESULT	3				Db	48 GTATPS 53				
Q9Q0H8		PRELIMINARY;	PRT;	80 AA.	RESULT	5	PRELIMINARY;	PRT;	92 AA.	
ID	Q9Q0H8				ID	Q38158				
AC	Q9Q0H8;				AC	Q38158;				
DT	01-DEC-2001	(TREMBLrel. 19, Created)			DT	01-JAN-1998	(TREMBLrel. 05, Created)			
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)			DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	GAG PROTEIN (FRAGMENT).				DE	GAG PROTEIN (FRAGMENT).				
GN	GAG.				GN	GAG.				
OS	Human immunodeficiency virus type 1.				OS	Human immunodeficiency virus type 1.				
OC	Viruses; Retroviridae; Lentivirus.				OC	Viruses; Retrovirus; Lentivirus.				
NCBL-TaxID=11676;					OX	NCBL-TaxID=11676;				
RN					RN					
RP	SEQUENCE FROM N.A.				RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21248068; PubMed=11350662;				RC	SPRINKLE-PATENT 4.				
RA	Kaufmann G.R., Suzuki K., Cunningham P., Mukaiide M., Kondo M., Imai M., Saunders J., Cooper D.A.;				RC	MEDLINE=9744676; PubMed=9261388;				
RT	"Impact of HIV Type 1 Protease, Reverse Transcriptase, Cleavage Site, and p6 Mutations on the Virological Response to Quadruplet Therapy with Saquinavir, Ritonavir, and Two Nucleoside Analogs."				RA	Zhang Y.M., Imamichi H., Imamichi N.P.; Vasudevachai M.B.; Salzman N.P.;				
RT	AIDS Res. Hum. Retroviruses 17:487-497(2001).				RT	"Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites."				
RT	EMBL; AF323213; AAK6674.1; -.				RT	J. Virol. 71:6662-6670(1997).				
DR	NON_TER 1 1				RL	DR				
SEQUENCE	80 AA; 8942 MW; 72C5C3148E51E65D CRC64;				DR	EMBL; AF024554; AAB83033.1; -.				
Qy	1 grtxxps 6				DR	HSSP; P05888; 1AAF.				
Db	48 GTATPS 53				DR	InterPro; IPR001878; Znf_CCHC.				
Query Match	91.7%; Score 22; DB 15; Length 92;				DR	PFam; PF000088; zf_CCHC; 1.				
Best Local Similarity	91.7%; Score 22; DB 15; Length 92;				DR	SMART; SM00434; ZnF_C2HC; 2.				
Matches	4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				KW	Zinc-finger				
Qy	1 grtxxps 6				FT	NON_TER 1				
Db	48 GTATPS 53				FT	NON_TER 92 AA; 10403 MW; 27E1160C29F61EA CRC64;				
Query Match	91.7%; Score 22; DB 15; Length 92;				SQ	SEQUENCE 92 AA; 10403 MW; 27E1160C29F61EA CRC64;				

Best Local Similarity	66.7%	Pred.	No. 1.3e+02;		
Matches	4;	Conservative	0;	Mismatches	2;
Qy	1	9ttxps	6		
Db	72	GTATPS	77		
RESULT	6				
O38175					
ID	O38175	PRELIMINARY;			
AC	O38175;				
DT	01-JAN-1998	(TREMBLrel. 05, Created)			
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	GAG PROTEIN (FRAGMENT).				
GN	GAG.				
OS	Human immunodeficiency virus type 1.				
OC	Viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxID:11676;				
RN	[1]				
SEQUENCE	FROM N.A.				
STRAIN	=PATIENT 4;				
RX	MEDLINE=77404676; PubMed=9261388;				
RA	Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,				
RA	Vasudevachari M.B., Salzman N.P.;				
RT	"Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites."				
RT	J. Virol. 71:6662-6670(1997).				
DR	EMBL; AF024071; AAB83850; 1.;				
DR	HSSP; P05888; 1AAF.				
DR	InterPro; IPR001878; Znf_CCHC.				
DR	PFAM; PF00098; zf-CCHC; 1.				
DR	SMART; SM00343; Znf_C2HC; 2.				
DR	Zinc-finger.				
FT	NON_TER 1	1			
FT	NON_TER 93	93			
FT	SEQUENCE 93 AA;	10402 MW:	0E9FE41DC4129C73 CRC64;		
Query Match	91.7%	Score 22;	DB 15;	Length 93;	
Best Local Similarity	66.7%	Pred.	No. 1.3e+02;		
Matches	4;	Conservative	0;	Mismatches	2;
Qy	1	9ttxps	6		
Db	73	GTATPS	78		
RESULT	7				
O38177					
ID	O38177	PRELIMINARY;			
AC	O38177;				
DT	01-JAN-1998	(TREMBLrel. 05, Created)			
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	GAG PROTEIN (FRAGMENT).				
GN	GAG.				
OS	Human immunodeficiency virus type 1.				
OC	Viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxID:11676;				
RN	[1]				
SEQUENCE	FROM N.A.				
STRAIN	=PATIENT 4.				
RX	MEDLINE=97404676; PubMed=9261388;				
RA	Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,				
RA	Vasudevachari M.B., Salzman N.P.;				
RT	"Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites."				
RT	J. Virol. 71:6662-6670(1997).				
DR	HSSP; P05888; 1AAF.				
DR	InterPro; IPR001878; Znf_CCHC.				

RX MEDLINE=97404676; PubMed=9261388;	DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;	DE GAG PROTEIN (FRAGMENT).
RA Human immunodeficiency virus type 1.	GN GAG.
RT "Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";	OS Viruses; Retroviridae; Lentivirus
RT J. Virol. 71:6662-6670(1997).;	OC
RT EMBL: AF024057; AAB83836.1; -.	OX NCBI_TAXID=11676;
DR HSSP; P05888; 1AAF.	RN [1]
DR IntePro; IPR001878; Znf_CCHC.	RP SEQUENCE FROM N.A.
DR Pfam; PF00098; zf-CCHC; 2.	RC STRAIN=PATIENT 3;
SMART; SM00343; Znf_C2HC; 2.	RX MEDLINE=97404676; PubMed=9261388;
DR Zinc-finger.	RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;
KW NON_TER 1 1	RA "Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";
FT NON_TER 96 96	RT J. Virol. 71:6662-6670(1997).;
SEQUENCE 96 AA; 10824 MW; 74B74DFE73E93935 CRC64;	RL DR EMBL; AF023926; AAB83860.1; -.
DR HSSP; P05888; 1AAF.	DR HSSP; P05888; 1AAF.
DR InterPro; IPR001878; Znf_CCHC.	DR InterPro; IPR001878; Znf_CCHC.
DR SMART; SM00343; Znf_C2HC; 2.	DR SMART; SM00343; Znf_C2HC; 2.
DR Zinc-finger.	KW Zinc-finger.
FT NON_TER 1 1	FT NON_TER 1 1
SEQUENCE 99 AA; 11140 MW; 92348443928FDCC0 CRC64;	SQ SEQUENCE 99 AA; 11140 MW; 92348443928FDCC0 CRC64;
Query Match 91.7%; Score 22; DB 15; Length 96;	Query Match 91.7%; Score 22; DB 15; Length 99;
Best Local Similarity 66.7%; Pred. No. 1.3e-02; Mismatches 2; Indels 0; Gaps 0;	Best Local Similarity 66.7%; Pred. No. 1.4e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 9ttxxps 6	Qy 1 9ttxxps 6
Db 76 GTTTTS 81	Db 79 GTTTTS 84
RESULT 10	RESULT 12
038178 ID 038178 PRELIMINARY; PRT; 96 AA.	038362 ID 038362 PRELIMINARY; PRT; 99 AA.
AC 038178; STRAIN=PATIENT 4; RC 01-JAN-1998 (TREMBLrel. 05, Created)	AC 038362; STRAIN=PATIENT 4; RC 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)	DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)	DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)	DE GAG PROTEIN (FRAGMENT).
DE GAG PROTEIN (FRAGMENT).	GN GAG.
GN Human immunodeficiency virus type 1.	OS Viruses; Retroviridae; Lentivirus
OS Viruses; Retroviridae; Lentivirus.	OC
OC NCBITAXID=11676; RN [1]	OX NCBI_TAXID=11676;
RP SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RC STRAIN=PATIENT 4; RX MEDLINE=97404676; PubMed=9261388;	RC STRAIN=PATIENT 7;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;	RX MEDLINE=97404676; PubMed=9261388;
RT "Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";	RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;
RT J. Virol. 71:6662-6670(1997).;	RA "Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";
RT EMBL: AF02407; AAB83853.1; -.	RT J. Virol. 71:6662-6670(1997).;
DR HSSP; P05888; 1AAF.	DR DR EMBL; AF02407; AAB83853.1; -.
DR IntePro; IPR001878; Znf_CCHC.	DR HSSP; P05888; 1AAF.
DR Pfam; PF00098; zf-CCHC; 2.	DR IntePro; IPR001878; Znf_CCHC.
SMART; SM00343; Znf_C2HC; 2.	DR SMART; SM00343; Znf_C2HC; 2.
DR Zinc-finger.	KW Zinc-finger.
KW NON_TER 1 1	FT NON_TER 1 1
SEQUENCE 96 AA; 10794 MW; 74B756FFC3E93935 CRC64;	SEQUENCE 99 AA; 11063 MW; 9818C494BED11470 CRC64;
Query Match 91.7%; Score 22; DB 15; Length 96;	Query Match 91.7%; Score 22; DB 15; Length 99;
Best Local Similarity 66.7%; Pred. No. 1.3e-02; Mismatches 2; Indels 0; Gaps 0;	Best Local Similarity 66.7%; Pred. No. 1.4e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 9ttxxps 6	Qy 1 9ttxxps 6
Db 76 GTATPS 81	Db 79 GTATPS 84
RESULT 11	RESULT 11
038030 ID 038030 PRELIMINARY; PRT; 99 AA.	038030 ID 038030 PRELIMINARY; PRT; 99 AA.
AC 038030; STRAIN=PATIENT 4; RC 01-JAN-1998 (TREMBLrel. 05, Created)	AC 038030; STRAIN=PATIENT 4; RC 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)	DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE GAG PROTEIN (FRAGMENT).	DE GAG PROTEIN (FRAGMENT).
GN GAG.	OS Viruses; Retroviridae; Lentivirus
OS Human immunodeficiency virus type 1.	OC
OC Viruses; Retroviridae; Lentivirus.	OX NCBI_TAXID=11676;
OC NCBITAXID=11676; RN [1]	RP SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.	RC STRAIN=PATIENT 7;
RC STRAIN=PATIENT 4; RX MEDLINE=97404676; PubMed=9261388;	RX MEDLINE=97404676; PubMed=9261388;
RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;	RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;
RT "Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";	RA "Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";
RT J. Virol. 71:6662-6670(1997).;	RT J. Virol. 71:6662-6670(1997).;
RT EMBL: AF02407; AAB83853.1; -.	RT DR EMBL; AF02407; AAB83853.1; -.
DR HSSP; P05888; 1AAF.	DR HSSP; P05888; 1AAF.
DR IntePro; IPR001878; Znf_CCHC.	DR IntePro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 2.	DR SMART; SM00343; Znf_C2HC; 2.
SMART; SM00343; Znf_C2HC; 2.	KW Zinc-finger.
DR Zinc-finger.	FT NON_TER 1 1
KW NON_TER 1 1	SEQUENCE 99 AA; 11063 MW; 9818C494BED11470 CRC64;
SEQUENCE 96 AA; 10794 MW; 74B756FFC3E93935 CRC64;	SQ SEQUENCE 99 AA; 11063 MW; 9818C494BED11470 CRC64;
Query Match 91.7%; Score 22; DB 15; Length 96;	Query Match 91.7%; Score 22; DB 15; Length 99;
Best Local Similarity 66.7%; Pred. No. 1.3e-02; Mismatches 2; Indels 0; Gaps 0;	Best Local Similarity 66.7%; Pred. No. 1.4e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 9ttxxps 6	Qy 1 9ttxxps 6
Db 76 GTATPS 81	Db 79 GTATPS 84

Db	79	GTATPS 84	SQ	SEQUENCE	99 AA;	10934 MW;	2808C494BEC79C6 CRC64;
RESULT 13							
ID	038371	PRELIMINARY;	PRT;	99 AA.			
AC	038371;						
DT	01-JAN-1998	(TREMBLrel.	05;	Created)			
DT	01-JAN-1998	(TREMBLrel.	05;	Last sequence update)			
DT	01-DEC-2001	(TREMBLrel.	19;	Last annotation update)			
DE		GAG PROTEIN (FRAGMENT).					
GN		GAG.					
OS		Human immunodeficiency virus type 1.					
OC		Viruses; Retroviroidea; Lentiviridae; Lentivirus.					
OX		NCBI_TaxID=11676;					
RN	[1]						
RP		SEQUENCE FROM N.A.					
RC		STRAIN=PATIENT 7;					
RX		MEDLINE=97404676; PubMed=9261388;					
RA		Zhang Y.M., Imanichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;					
RA		"Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";					
RT		J. Virol. 71:6662-6670(1997).					
RL		EMBL: AF024270; AAB83144.1;					
DR		HSSP; P05888; 1AAF.					
DR		InterPro; IPR01878; Znf_CCHC.					
DR		PFAM; PF00098; zf-CCNC; 2.					
DR		SMART; SM00343; Znf_C2HC; 2.					
KW		Zinc-finger.					
FT		NON_TER 99 99 1 1					
FT		NON_TER 99 99 1 1					
SQ		SEQUENCE 99 AA;	11092 MW;	9818C494BD1E1470 CRC64;			
Query Match		91.7%;	Score 22;	DB 15;	Length 99;		
Best Local Similarity		66.7%;	Pred. No. 1.4e+02;				
Matches	4;	Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;	
Qy	1	gttxps 6					
Db	79	GTATPS 84					
RESULT 14							
ID	038372	PRELIMINARY;	PRT;	99 AA.			
AC	038372;						
DT	01-JAN-1998	(TREMBLrel.	05;	Created)			
DT	01-JAN-1998	(TREMBLrel.	05;	Last sequence update)			
DT	01-DEC-2001	(TREMBLrel.	19;	Last annotation update)			
DE		GAG PROTEIN (FRAGMENT).					
GN		GAG.					
OS		Human immunodeficiency virus type 1.					
OC		Viruses; Retroviroidea; Lentiviridae; Lentivirus.					
OX		NCBI_TaxID=11676;					
RN	[1]						
RP		SEQUENCE FROM N.A.					
RC		STRAIN=PATIENT 7;					
RX		MEDLINE=97404676; PubMed=9261388;					
RA		Zhang Y.M., Imanichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;					
RA		"Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";					
RT		J. Virol. 71:6662-6670(1997).					
DR		EMBL: AF024271; AAB83145.1;					
DR		InterPro; IPR01878; Znf_CCHC.					
DR		PFAM; PF00098; zf-CCNC; 2.					
KW		Zinc-finger.					
FT		NON_TER 1 1					
FT		NON_TER 101 101					
SQ		SEQUENCE 101 AA;	11408 MW;	2EF1A3E5CF282854 CRC64;			
Query Match		91.7%;	Score 22;	DB 15;	Length 101;		
Best Local Similarity		66.7%;	Pred. No. 1.4e+02;				
Matches	4;	Conservative	0;	Mismatches 2;	Indels 2;	Gaps 0;	
Qy	1	gttxps 6					
Db	79	GTATPS 84					
RESULT 15							
ID	038113	PRELIMINARY;	PRT;	101 AA.			
AC	038113;						
DT	01-JAN-1998	(TREMBLrel.	05;	Created)			
DT	01-JAN-1998	(TREMBLrel.	05;	Last sequence update)			
DT	01-DEC-2001	(TREMBLrel.	19;	Last annotation update)			
DE		GAG PROTEIN (FRAGMENT).					
GN		GAG.					
OS		Human immunodeficiency virus type 1.					
OC		Viruses; Retroviroidea; Lentiviridae; Lentivirus.					
OX		NCBI_TaxID=11676;					
RN	[1]						
RP		SEQUENCE FROM N.A.					
RC		STRAIN=PATIENT 7;					
RX		MEDLINE=97404676; PubMed=9261388;					
RA		Zhang Y.M., Imanichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;					
RA		"Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";					
RT		J. Virol. 71:6662-6670(1997).					
DR		EMBL: AF024271; AAB83145.1;					
DR		InterPro; IPR01878; Znf_CCHC.					
DR		PFAM; PF00098; zf-CCNC; 2.					
KW		Zinc-finger.					
FT		NON_TER 1 1					
FT		NON_TER 101 101					
SQ		SEQUENCE 101 AA;	11408 MW;	2EF1A3E5CF282854 CRC64;			
Query Match		91.7%;	Score 22;	DB 15;	Length 101;		
Best Local Similarity		66.7%;	Pred. No. 1.4e+02;				
Matches	4;	Conservative	0;	Mismatches 2;	Indels 2;	Gaps 0;	
Qy	1	gttxps 6					
Db	79	GTATPS 84					
RESULT 16							
ID	038128	PRELIMINARY;	PRT;	101 AA.			
AC	038128;						
DT	01-JAN-1998	(TREMBLrel.	05;	Created)			
DT	01-JAN-1998	(TREMBLrel.	05;	Last sequence update)			
DT	01-DEC-2001	(TREMBLrel.	19;	Last annotation update)			
DE		GAG PROTEIN (FRAGMENT).					
GN		GAG.					
OS		Human immunodeficiency virus type 1.					
OC		Viruses; Retroviroidea; Lentiviridae; Lentivirus.					
OX		NCBI_TaxID=11676;					
RN	[1]						
RP		SEQUENCE FROM N.A.					
RC		STRAIN=PATIENT 7;					
RX		MEDLINE=97404676; PubMed=9261388;					
RA		Zhang Y.M., Imanichi H., Imamichi T., Lane H.C., Falloon J., Vasudevachari M.B., Salzman N.P.;					
RA		"Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.";					
RT		J. Virol. 71:6662-6670(1997).					
DR		EMBL: AF024271; AAB83145.1;					
DR		InterPro; IPR01878; Znf_CCHC.					
DR		PFAM; PF00098; zf-CCNC; 2.					
KW		Zinc-finger.					
FT		NON_TER 1 1					
FT		NON_TER 101 101					
SQ		SEQUENCE 101 AA;	11408 MW;	2EF1A3E5CF282854 CRC64;			

RESULT 17
 O38130 ID 038130 PRELIMINARY; PRT; 101 AA.
 AC 018130; 0; 0;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 Human immunodeficiency virus type 1.
 OC Viruses; Retroviridae; Lentivirus.
 NCBI_TAXID=11676;
 OX RN [1]
 RP SEQUENCE FROM N.A.
 STRAIN=PATIENT 4.
 PC MEDLINE=7404676; PubMed=9261388;
 RX Zhang Y.M., Imamichi H., Inamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 the protease gene and in its Gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024024; AAB83803.1;
 HSSP; P05888; 1AAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON-TER 1 101
 FT NON-TER 101 101
 SQ SEQUENCE 101 AA; 11295 MW; EDD5B3278BF1D5FA CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 81 GTRTPS 86

RESULT 18
 O38133 ID 038133 PRELIMINARY; PRT; 101 AA.
 AC 038133; 0; 0;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 Human immunodeficiency virus type 1.
 OC Viruses; Retroviridae; Lentivirus.
 NCBI_TAXID=11676;

Query Match 91.7%; Score 22; DB 15; Length 101;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 81 GTRTPS 86

RESULT 19
 O38138 ID 038138 PRELIMINARY; PRT; 101 AA.
 AC 038138;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 Human immunodeficiency virus type 1.
 OC Viruses; Retroviroidea; Lentivirus.
 OX RN [1]
 RP SEQUENCE FROM N.A.
 STRAIN=PATIENT 4.
 PC MEDLINE=97404676; PubMed=9261388;
 RX Zhang Y.M., Imamichi H., Inamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 the protease gene and in its Gag substrate cleavage sites.";
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; AF024024; AAB83805.1;
 HSSP; P05888; 1AAF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR SMART; SM00343; Znf_C2HC; 2.
 KW Zinc-finger.
 FT NON-TER 1 101
 FT NON-TER 101 101
 SQ SEQUENCE 101 AA; 11408 MW; F7B36D60472BDE0 CRC64;

Query Match 91.7%; Score 22; DB 15; Length 101;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gtxxps 6
 Db 81 GTRTPS 86

RESULT 20
 O38141 ID 038141 PRELIMINARY; PRT; 101 AA.
 AC 038141; 0; 0;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 Human immunodeficiency virus type 1.
 OC Viruses; Retroviridae; Lentivirus.

Q38141; NCBI_TAXID=11676; RN [1]; RP SEQUENCE FROM N.A.	PRELIMINARY; PRT; 101 AA.	QY 1 gtxxps 6 DB 15 GTSSPS 20
RA Human immunodeficiency virus type 1. Viruses; Retrovirus; Lentiviridae; Lentivirus.		
RA Vasudevachari M.B.; Salzman N.P.; Zhang Y.M.; Imamichi H.; Lane H.C.; Falloon J.; RT "Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites."; RT J. Virol. 71:6662-6670(1997). DR AF024037; AAB83816.1; DR HSSP; P05888; IAAF. DR InterPro; IPR001878; Znf_CCHC. DR PF00098; zf_CCHC; 2. DR SMART; SM00343; Znf_C2HC; 2. KW Zinc-finger. FT NON_TER 1 1 FT NON_TER 101 101 SQ SEQUENCE 101 AA; 11379 MW; F7BB323A5781D5FA CRC64;		
Query Match 91.7%; Score 22; DB 15; Length 101; Best Local Similarity 66.7%; Pred. No. 1.4e+02; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY 1 gtxxps 6 DB 81 GTATPS 86		
RESULT 21		
Q9YA39; NCBI_TAXID=56636; RN [1]; RP SEQUENCE FROM N.A.	PRELIMINARY; PRT; 105 AA.	QY 1 gtxxps 6 DB 91 GTTPS 96
RA Kawarabayasi Y.; Hino Y.; Horikawa H.; Yamazaki S.; Haikawa Y.; Jin-no K.; Takahashi M.; Sekine M.; Baba S.-I.; Anzai A.; Kosugi H.; RA Hosoyama A.; Fukui S.; Nagai Y.; Nishijima K.; Nakazawa H.; RA Takamiya M.; Masuda S.; Funahashi T.; Tanaka T.; Kudo H.; Yamazaki J.; Kushiida N.; Oguchi A.; Aoki K.-I.; Kubota K.; RA Nakamura Y.; Nomura N.; Sako Y.; Kikuchi H.; RT "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1."; RL DNA Res. 6:83-101(1999). DR EMBL; AP00063; BAA81111.1; KW Hypothetical protein; Complete proteome. SQ SEQUENCE 105 AA; 11262 MW; B25D94F6CED05D87 CRC64;	PRELIMINARY; PRT; 105 AA.	PRELIMINARY; PRT; 111 AA.
Query Match 91.7%; Score 22; DB 17; Length 105; Best Local Similarity 66.7%; Pred. No. 1.5e+02; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Q9YA39; NCBI_TAXID=11676; RN [1]; RP SEQUENCE FROM N.A.	PRELIMINARY; PRT; 105 AA.	QY 1 gtxxps 6 DB 91 GTTPS 96
RA Zhang Y.M.; Imamichi H.; Imamichi T.; Lane H.C.; Falloon J.; RT "Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites."; RT J. Virol. 71:6662-6670(1997). DR AF024037; AAB83816.1; DR HSSP; P05888; IAAF. DR InterPro; IPR001878; Znf_CCHC. DR PF00098; zf_CCHC; 2. DR SMART; SM00343; Znf_C2HC; 2.	PRELIMINARY; PRT; 111 AA.	PRELIMINARY; PRT; 111 AA.
Query Match 91.7%; Score 22; DB 15; Length 111; Best Local Similarity 66.7%; Pred. No. 1.5e+02; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Q9YA39; NCBI_TAXID=11676; RN [1]; RP SEQUENCE FROM N.A.	PRELIMINARY; PRT; 105 AA.	QY 1 gtxxps 6 DB 91 GTTPS 96
RA Kawarabayasi Y.; Hino Y.; Horikawa H.; Yamazaki S.; Haikawa Y.; Jin-no K.; Takahashi M.; Sekine M.; Baba S.-I.; Anzai A.; Kosugi H.; RA Hosoyama A.; Fukui S.; Nagai Y.; Nishijima K.; Nakazawa H.; RA Takamiya M.; Masuda S.; Funahashi T.; Tanaka T.; Kudo H.; Yamazaki J.; Kushiida N.; Oguchi A.; Aoki K.-I.; Kubota K.; RA Nakamura Y.; Nomura N.; Sako Y.; Kikuchi H.; RT "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1."; RL DNA Res. 6:83-101(1999). DR EMBL; AP00063; BAA81111.1; KW Hypothetical protein; Complete proteome. SQ SEQUENCE 105 AA; 11262 MW; B25D94F6CED05D87 CRC64;	PRELIMINARY; PRT; 111 AA.	PRELIMINARY; PRT; 111 AA.
Query Match 91.7%; Score 22; DB 17; Length 105; Best Local Similarity 66.7%; Pred. No. 1.5e+02; Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		

KW Zinc-finger.
 FT NON_TER 1 1
 FT NON_TER 111 111
 SQ SEQUENCE 111 AA; FC3BEEFB25A1B7EDB CRC64;

Query Match 24
 Best Local Similarity 91.7%; Score 22; DB 15; Length 111;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 91 GRTTPS 96

RESULT 038169 PRELIMINARY; PRT; 111 AA.
 ID 038169; (TREMBLrel. 05, Created)
 AC 038169; (TREMBLrel. 05, Last sequence update)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PATIENT 4;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,
 RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 the protease gene and in its Gag substrate cleavage sites."
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; IPRO01878; 1AAE.
 DR HSSP; P05888; 1AAE.
 DR InterPro; IPRO01878; 2nf_CCHC.
 DR SMART; SM00343; 2nf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 SQ SEQUENCE 111 AA; 12684 MW; 39C82CAA0BADFDDC CRC64;

Query Match 25
 Best Local Similarity 91.7%; Score 22; DB 15; Length 111;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 91 GRTTPS 96

RESULT 038171 PRELIMINARY; PRT; 111 AA.
 ID 038171; (TREMBLrel. 05, Created)
 AC 038171; (TREMBLrel. 05, Last sequence update)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PATIENT 4;
 RX MEDLINE=97404676; PubMed=9261388;
 RA Zhang Y.M., Imamichi H., Imamichi T., Lane H.C., Falloon J.,

RA Vasudevachari M.B., Salzman N.P.;
 RT "Drug resistance during indinavir therapy is caused by mutations in
 the protease gene and in its Gag substrate cleavage sites."
 RL J. Virol. 71:6662-6670(1997).
 DR EMBL; IPRO01878; 2nf_CCHC.
 DR HSSP; P05888; 1AAE.
 DR InterPro; IPRO01878; 2nf_CCHC.
 DR SMART; SM00343; 2nf_C2HC; 2.
 KW Zinc-finger.
 FT NON_TER 1 1
 SQ SEQUENCE 112 AA; 12895 MW; C97021F2A4407CA7 CRC64;

Query Match 26
 Best Local Similarity 91.7%; Score 22; DB 15; Length 112;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 91 GRTTPS 96

RESULT 038172 PRELIMINARY; PRT; 135 AA.
 ID 038172; (TREMBLrel. 05, Created)
 AC 038172; (TREMBLrel. 05, Last sequence update)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DE GAG PROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;

Query Match 27
 Best Local Similarity 91.7%; Score 22; DB 15; Length 112;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 94 GRTTPS 99

Query Match 28
 Best Local Similarity 91.7%; Score 22; DB 15; Length 112;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gtxxps 6
 Db 94 GRTTPS 99

DE	ORF 5' OF FIBROBLAST GROWTH FACTOR RECEPTOR 1 FGFR-1.	Best Local Similarity 66.7%; Pred. No. 1.9e+02;
GN	FGFR1.	Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
MU	Mus sp.	
OS		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OC		
NCBI_TAXID	10095;	
OX		
[1]	RN	
RP	SEQUENCE FROM N.A. MEDLINE=95100936; PubMed=7802632;	
RX	RA Harada T., Saito H., Kohlbara H., Kurebayashi S., Kasayama S., Terakawa N., Kishimoto T., Saito B.,	
RA	RT "Murine fibroblast growth factor receptor 1 gene generates multiple messenger RNAs containing two open reading frames via alternative splicing."; Biochem. Biophys. Res. Commun. 205:1057-1063(1994).	
RT	RL EMBL: S74765; AB32844.2; -.	
DR	MGD; MGI:95522; Fgf1.	
KW	Receptor.	
SEQUENCE	135 AA; 14280 MW; 03906AD6DC40880F CRC64;	
Q9D5C6	RESULT 29	
ID Q9D5C6	PRELIMINARY;	PRT; 156 AA.
AC Q9D5C6;	ID Q9HBNT;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9HBNT;	PRT; 156 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
RT	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
RL	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
DE	DE HYPOTHETICAL 16.2 KDA PROTEIN.	
OS	OS Homo sapiens (Human).	
OC	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=9606;	
OX	OX RN [1]	
RP	RP SEQUENCE FROM N.A.	
RA	RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Li H.N., Yu Y.,	
RA	RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,	
RA	RA Yu J., Han L.H.;	
RT	RT "Novel Human cDNA clones with function of inhibiting cancer cell growth."	
RL	RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.	
DR	DR AF218021; AAG17263.1; -.	
KW	KW HYPOTHETICAL PROTEIN.	
SQ	SQ SEQUENCE 156 AA; 16178 MW; 183180BEC8F93E63 CRC64;	
Q9D5C6	RESULT 28	
ID Q9D5C6	PRELIMINARY;	PRT; 142 AA.
AC Q9D5C6;	ID Q9NWG3;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9NWG3;	PRT; 157 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DE 4930465A12RIK PROTEIN.	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
GN 4930465A12RIK.	DE HYPOTHETICAL 17.4 KDA PROTEIN.	
OS Mus musculus (Mouse).	OS Homo sapiens (Human).	
OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Murinae; Muridae; Mus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=10090;	
RN [1]	RN RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC STRAIN=C57BL/6J; TISSUE=TESTIS;	
RX	RX MEDLINE=21085680; PubMed=11217851;	
RA	RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,	
RA	RA Kondo H., Konno H., Adachi J., Fukuda S.,	
RA	RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,	
RA	RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,	
RA	RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,	
RA	RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,	
RA	RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,	
RA	RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,	
RA	RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,	
RA	RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,	
RA	RA Brownstein M.J., Bult C., Flechner C., Gariboldi M.,	
RA	RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,	
RA	RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,	
RA	RA Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N.,	
RA	RA Sasaki H., Sato K., Schoenbach C., Seye T., Shibata Y., Storch K.-F.,	
RA	RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,	
RA	RA Watanabe-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,	
RA	RA Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection."	
RN	RN EMBL; AK015502; BAB29872.1; -.	
DR	DR MGD; MGI:1922239; 4930465A12RIK.	
SEQUENCE	SEQUENCE 142 AA; 15397 MW; AF4B8868A17B745E CRC64;	
Q9	Query Match 91.7%; Score 22; DB 11; Length 142;	
Q9D5C6	RESULT 27	
ID Q9D5C6	PRELIMINARY;	PRT; 142 AA.
AC Q9D5C6;	ID Q9NWG3;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9NWG3;	PRT; 157 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DE 4930465A12RIK PROTEIN.	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
GN 4930465A12RIK.	DE HYPOTHETICAL 17.4 KDA PROTEIN.	
OS Mus musculus (Mouse).	OS Homo sapiens (Human).	
OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Murinae; Muridae; Mus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=10090;	
RN [1]	RN RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC TISSUE=EMRYO;	
RX	RX Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,	
RA	RA Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,	
RA	RA Tanase T., Nomura Y., Togiya S., Komai F., Takeuchi K.,	
RA	RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,	
RA	RA Wakanatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;	
RA	RA "NEO human cDNA sequencing project."	
RT	RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
DR	DR EMBL; AK000905; BAA1418.1; -.	
SEQUENCE	SEQUENCE 157 AA; 17352 MW;	
Q9	Query Match 91.7%; Score 22; DB 4; Length 157;	
Q9D5C6	RESULT 26	
ID Q9D5C6	PRELIMINARY;	PRT; 142 AA.
AC Q9D5C6;	ID Q9NWG3;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9NWG3;	PRT; 157 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DE 4930465A12RIK PROTEIN.	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
GN 4930465A12RIK.	DE HYPOTHETICAL 17.4 KDA PROTEIN.	
OS Mus musculus (Mouse).	OS Homo sapiens (Human).	
OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Murinae; Muridae; Mus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=10090;	
RN [1]	RN RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC TISSUE=EMRYO;	
RX	RX Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,	
RA	RA Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,	
RA	RA Tanase T., Nomura Y., Togiya S., Komai F., Takeuchi K.,	
RA	RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,	
RA	RA Wakanatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;	
RA	RA "NEO human cDNA sequencing project."	
RT	RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
DR	DR EMBL; AK000905; BAA1418.1; -.	
SEQUENCE	SEQUENCE 157 AA; 17352 MW;	
Q9	Query Match 91.7%; Score 22; DB 4; Length 157;	
Q9D5C6	RESULT 25	
ID Q9D5C6	PRELIMINARY;	PRT; 142 AA.
AC Q9D5C6;	ID Q9NWG3;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9NWG3;	PRT; 157 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DE 4930465A12RIK PROTEIN.	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
GN 4930465A12RIK.	DE HYPOTHETICAL 17.4 KDA PROTEIN.	
OS Mus musculus (Mouse).	OS Homo sapiens (Human).	
OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Murinae; Muridae; Mus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=10090;	
RN [1]	RN RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC TISSUE=EMRYO;	
RX	RX Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,	
RA	RA Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,	
RA	RA Tanase T., Nomura Y., Togiya S., Komai F., Takeuchi K.,	
RA	RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,	
RA	RA Wakanatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;	
RA	RA "NEO human cDNA sequencing project."	
RT	RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
DR	DR EMBL; AK000905; BAA1418.1; -.	
SEQUENCE	SEQUENCE 157 AA; 17352 MW;	
Q9	Query Match 91.7%; Score 22; DB 4; Length 157;	
Q9D5C6	RESULT 24	
ID Q9D5C6	PRELIMINARY;	PRT; 142 AA.
AC Q9D5C6;	ID Q9NWG3;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9NWG3;	PRT; 157 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DE 4930465A12RIK PROTEIN.	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
GN 4930465A12RIK.	DE HYPOTHETICAL 17.4 KDA PROTEIN.	
OS Mus musculus (Mouse).	OS Homo sapiens (Human).	
OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Murinae; Muridae; Mus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=10090;	
RN [1]	RN RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC TISSUE=EMRYO;	
RX	RX Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,	
RA	RA Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,	
RA	RA Tanase T., Nomura Y., Togiya S., Komai F., Takeuchi K.,	
RA	RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,	
RA	RA Wakanatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;	
RA	RA "NEO human cDNA sequencing project."	
RT	RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
DR	DR EMBL; AK000905; BAA1418.1; -.	
SEQUENCE	SEQUENCE 157 AA; 17352 MW;	
Q9	Query Match 91.7%; Score 22; DB 4; Length 157;	
Q9D5C6	RESULT 23	
ID Q9D5C6	PRELIMINARY;	PRT; 142 AA.
AC Q9D5C6;	ID Q9NWG3;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9NWG3;	PRT; 157 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DE 4930465A12RIK PROTEIN.	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
GN 4930465A12RIK.	DE HYPOTHETICAL 17.4 KDA PROTEIN.	
OS Mus musculus (Mouse).	OS Homo sapiens (Human).	
OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Murinae; Muridae; Mus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=10090;	
RN [1]	RN RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC TISSUE=EMRYO;	
RX	RX Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,	
RA	RA Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,	
RA	RA Tanase T., Nomura Y., Togiya S., Komai F., Takeuchi K.,	
RA	RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,	
RA	RA Wakanatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;	
RA	RA "NEO human cDNA sequencing project."	
RT	RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
DR	DR EMBL; AK000905; BAA1418.1; -.	
SEQUENCE	SEQUENCE 157 AA; 17352 MW;	
Q9	Query Match 91.7%; Score 22; DB 4; Length 157;	
Q9D5C6	RESULT 22	
ID Q9D5C6	PRELIMINARY;	PRT; 142 AA.
AC Q9D5C6;	ID Q9NWG3;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9NWG3;	PRT; 157 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DE 4930465A12RIK PROTEIN.	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
GN 4930465A12RIK.	DE HYPOTHETICAL 17.4 KDA PROTEIN.	
OS Mus musculus (Mouse).	OS Homo sapiens (Human).	
OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Murinae; Muridae; Mus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=10090;	
RN [1]	RN RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC TISSUE=EMRYO;	
RX	RX Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,	
RA	RA Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,	
RA	RA Tanase T., Nomura Y., Togiya S., Komai F., Takeuchi K.,	
RA	RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,	
RA	RA Wakanatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;	
RA	RA "NEO human cDNA sequencing project."	
RT	RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
DR	DR EMBL; AK000905; BAA1418.1; -.	
SEQUENCE	SEQUENCE 157 AA; 17352 MW;	
Q9	Query Match 91.7%; Score 22; DB 4; Length 157;	
Q9D5C6	RESULT 21	
ID Q9D5C6	PRELIMINARY;	PRT; 142 AA.
AC Q9D5C6;	ID Q9NWG3;	PRELIMINARY;
DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9NWG3;	PRT; 157 AA.
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DE 4930465A12RIK PROTEIN.	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
GN 4930465A12RIK.	DE HYPOTHETICAL 17.4 KDA PROTEIN.	
OS Mus musculus (Mouse).	OS Homo sapiens (Human).	
OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Murinae; Muridae; Mus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	OC NCBI_TAXID=10090;	
RN [1]	RN RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC TISSUE=EMRYO;	
RX	RX Isogai T., Ota T., Hay	

RESULT	3	04/1995	PRELIMINARY;	PRT;	177 AA.
Q9VSQ7	PRELIMINARY;	PRT;	175 AA.		
AC	04/1995;				
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)				
Q9VSQ7;					
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)				
DE	CG13309 PROTEIN.				
GN	CG13309.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Bukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydriidae; Drosophilidae; Drosophila.				
OX	NCBI_TAXID=2227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SEQUENCE IN BERKELEY;				
RX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celikin S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Richard S., Ashburner M., Henderson S.N.,				
RA	Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	Sutton G.G., Worms J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazquez R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.I.G.,				
RA	Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Borkov M.R., Bouck J., Brokstone P., Brotlier P.,				
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	De Pablos B., Deicher A., Deng Z., May S.A.D., Dietz S.M.,				
RA	Dodson K.J., Doup L.E., Dowres M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durbin K.J., Evangelista C.C., Ferreria S., Fleischmann W.,				
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gompel J., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J.J., Hernandez J.R., Houck J.,				
RA	Hostin D., Houston K.A., Howland T.J., Ke Z., Kenison J.A., Ketchum K.A.,				
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., KulP D., Lai Z.,				
RA	Lasko P., Lei J., Levitus A., Li J., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy L., Muzyk D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacieb J.M.,				
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,				
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,				
RA	Svirskas R., Tector C., Turner E., Venter J.C., Venter J.C.,				
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,				
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,				
RA	Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,				
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,				
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,				
RA	"The genome sequence of Drosophila melanogaster."				
RL	Science 287:2185-2195 (2000).				
DR	EMBL; AE003554; AAF50359.1; -.				
DR	FlyBase; FBgn0059393; CG13309.				
DR	FlyBase; FBgn0059393; CG13309.				
SEQUENCE	175 AA; 18038 MW; A094BA899ADFA1C7 CRC64;				
Qy	1 gtxxps 6				
SQ	Query Match 91.7%; Score 22; DB 5; Length 175;				
	Best Local Similarity 66.7%; Pred. No. 2.3e+02;				
	Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
Db	98 GSTS999 103				

Query Match	91.7%	Score 22;	DB 2 ² ;	Length 178;
Best Local Similarity	66.7%	Pred. No.	2.3e+02;	
Matches 4;	Conservative 7/8;	0;	Mismatches 2;	Indels 0;
Y	1 gtxrps 6			
Y	1			
Y	2 GTrSFS 7			
RESULT	34			
Q9VSQ6	PRELIMINARY;	PRT:	182 AA.	
Q9VSQ6;				
01-MAY-2000	(TREMBLrel. 13, Created)			
01-MAY-2000	(TREMBLrel. 13, Last sequence update)			
01-MAY-2000	(TREMBLrel. 13, Last annotation update)			
CG13308 PROTEIN.				
CG13308.				
Drosophila melanogaster (Fruit fly).				
Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
Ephydriidae; Drosophilidae; Drosophila.				
NCBI-TAXID=7221;				
[1] IN				
SEQUENCE FROM N.A.				
STRAIN=BERKELEY;				
MEDLINE=20196006; PubMed=10731132;				
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandao R.C., Rogers Y.-H.C., Blazquez R.G., Champe M., Pfeiffer B., Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G., Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch G., Baldwin B.M., Basu A., Boxendale J., Bayraktaroglu L., Beasley S., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Butler K.C., Busam D.A., Butler J., Cadieu E., Center A., Chandru Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Doddson K., Doup L.E., Downes M., Dugan-Rocha S., Dunker B.C., Du Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Gloede A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J., J.R., Houck J.J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalodis F., Karpen G.H., Ke Z., Kennison J.A., Keichum Kimmel B.E., Kodira C.D., Kraft C., Kravitz S.N., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattel Kiamos I., McInnis T.C., McPhereson D., Merkulov G., Malsinha N.V., Nobary C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M., Palazzolo M., Peltzman G.S., Pan S., Pollard J., Purvi V., Reese N., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng X.H., Zheng F.N., Zhong W., Zhou X., Zhu S., Zhu X., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., "The genome sequence of Drosophila melanogaster."; EMBL; AE003554; AAFF0360.1; -;				
SEQUENCE 182 AA;	18808 MW;	64B9775C44A3AEE2 CRC64;		
SEQUENCE 182 AA;	18808 MW;	64B9775C44A3AEE2 CRC64;		
Query Match	91.7%	Score 22;	DB 5;	Length 182;
Best Local Similarity	66.7%	Pred. No.	2.4e+02;	
Matches 4;	Conservative 4;	0;	Mismatches 2;	Indels 0;

Db	134	GTAPS 139	Qy	1 gtxxps 6	Db	206 GTATPS 211
RESULT 39						
Q9JLHS	Q9JLHS	PRELIMINARY;	PRT;	218 AA.	Q9KXV0	PRELIMINARY;
ID	Q9JLHS;				ID	PRT; 233 AA.
AC	Q9JLHS;				Q9KXV0;	
DT	01-OCT-2000 (TREMBLrel. 15, Created)				AC	(TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)				DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
CDKS ACTIVATOR BINDING PROTEIN.					DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE	Rattus norvegicus (Rat).				DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				DE	PUTATIVE SECRETED PROTEIN.
OC					GN	SCD95A_24.
OX	NCBI_TaxID:10116;				OS	Streptomyces coelicolor.
RN	[1]				OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetidae; Streptomyces.
SEQUENCE FROM N.A.					OC	Actinomycetales; Streptomyceae; Streptomyctaceae; Streptomyces.
RP	SEQUENCE FROM N.A.; PubMed=10721722;				NCBI_TaxID:1902;	
RX					RN	[1]
RA	Ching Y.P.; Qi Z.; Wang J.H.;				RP	SEQUENCE FROM N.A.
RT	*Cloning of three novel neuronal Cdk5 activator binding proteins in				RC	STRATN=A3(2);
RL	Gene 242:285-294(2000);				RA	Seeger K.J.; Harris D.;
DR	EMBL: AF177478; AAF60224; 1; -;				RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
SQ	SEQUENCE 218 AA; 24023 MW; 2C6750AA0A0D0EDA CRC64;				RN	[12]
Query Match						
Best Local Similarity	91.7%	Score 22;	DB 11;	Length 218;	RP	SEQUENCE FROM N.A.
Matches 4;	Conservative 66.7%;	Pred. No. 2.8e+02;	Indels 0;	Gaps 0;	RC	STRATN=A3(2);
Qy	1 gtxxps 6				RC	STRATN=A3(2);
Db	209 GTSSPS 214				RC	STRATN=A3(2);
RESULT 40						
Q9RDK2	Q9RDK2	PRELIMINARY;	PRT;	228 AA.	RX	MEDLINE=97000351; PubMed=8843436;
ID	Q9RDK2;				RA	Pedenbach M.; Kieser H.M.; Denapaitte D.; Eichner A.; Cullum J.,
AC	Q9RDK2;				RA	Kinashi H.; Hopwood D.A.;
DT	01-MAY-2000 (TREMBLrel. 13, Created)				RT	"A set of ordered cosmids and a detailed genetic and physical map for
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
DE	PUTATIVE MEMBRANE PROTEIN.				RL	Mol. Microbiol. 21:77-96(1996).
GN	SCC123_22C.				DR	EMBL: AL351432; CAB94051.1; -;
OS	Streptomyces coelicolor.				SQ	SEQUENCE 233 AA; 23071 MW; D3E0BEE9A3556083D CRC64;
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetaceae; Streptomyces.				RP	SEQUENCE FROM N.A.
OX	NCBI_TaxID:1902;				RC	STRATN=A3(2);
RN	[1]				RC	STRATN=A3(2);
RP	SEQUENCE FROM N.A.				RC	STRATN=A3(2);
RA	Brown S.P.; Harris D.;				RC	STRATN=A3(2);
RA	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				RC	STRATN=A3(2);
RL	[2]				RC	STRATN=A3(2);
RN	RP	SEQUENCE FROM N.A.;			RC	STRATN=A3(2);
RC	STRATN=A3(2);				RC	STRATN=A3(2);
RA	Thomson N.R.; Parkhill J.; Barrell B.G.; Rajandream M.A.;				RC	STRATN=A3(2);
RA	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				RC	STRATN=A3(2);
RL	[3]				RC	STRATN=A3(2);
RP	SEQUENCE FROM N.A.				RC	STRATN=A3(2);
RC	STRATN=A3(2);				RC	STRATN=A3(2);
RA	Medline=97000351; PubMed=8843436;				RC	STRATN=A3(2);
RA	Redenbach M.; Kieser H.M.; Denapaitte D.; Eichner A.; Cullum J.,				RC	STRATN=A3(2);
RA	Kinashi H.; Hopwood D.A.;				RC	STRATN=A3(2);
RT	"A set of ordered cosmids and a detailed genetic and physical map for				RC	STRATN=A3(2);
RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.";				RC	STRATN=A3(2);
RL	Mol. Microbiol. 21:77-96(1996).				RC	STRATN=A3(2);
DR	EMBL: AL351518; CAB66262.1; -;				RC	STRATN=A3(2);
SQ	SEQUENCE 228 AA; 24246 MW; 6217C85671050B30 CRC64;				RC	STRATN=A3(2);
Query Match						
Best Local Similarity	91.7%	Score 22;	DB 2;	Length 228;	Qy	1 gtxxps 6
Matches 4;	Conservative 66.7%;	Pred. No. 2.9e+02;	Indels 2;	Gaps 0;	Db	52 GTATPS 57
Qy	1 gtxxps 6				Qy	1 gtxxps 6
Db	206 GTATPS 211				Db	52 GTATPS 57
RESULT 42						
P79289	P79289	PRELIMINARY;	PRT;	237 AA.	Q9KXV0	PRELIMINARY;
ID	P79289;				ID	P79289;
AC	Q9KXV0;				AC	Q9KXV0;
DT	01-MAY-1997 (TREMBLrel. 03, Created)				DT	01-MAY-1997 (TREMBLrel. 03, Created)
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)				DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)
SP1 TRANSCRIPTION FACTOR (FRAGMENT).					DE	SP1 TRANSCRIPTION FACTOR (FRAGMENT).
OS	Sus scrofa (Pig).				OS	Sus scrofa (Pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Butcheria; Cetartiodactyla; Suina; Suidae; Sus.				OC	Mammalia; Butcheria; Cetartiodactyla; Suina; Suidae; Sus.
OX	NCBI_TaxID:9823;				OX	NCBI_TaxID:9823;
RN	SEQUENCE FROM N.A.				RN	SEQUENCE FROM N.A.
RP	TISSUE-ENDOMETRUM;				RP	TISSUE-ENDOMETRUM;
RC					RC	
RA	Medline=20150569; PubMed=10687861;				RA	Medline=20150569; PubMed=10687861;
RA	Silman R.C.M.; Zhang X.L.; Zhang D.; Wang Y.; Michel F.J.,				RA	Silman R.C.M.; Zhang X.L.; Zhang D.; Wang Y.; Michel F.J.,
RA	Slammon F.A.;				RA	Slammon F.A.;
RT	"Expression and regulatory function of the transcription factor Sp1 in				RT	"Expression and regulatory function of the transcription factor Sp1 in
RT	the uterine endometrium at early pregnancy: implications for				RT	the uterine endometrium at early pregnancy: implications for
RT	epithelial phenotype."				RT	epithelial phenotype."
RL	Mol. Cell. Endocrinol. 159:159-170(2000).				RL	Mol. Cell. Endocrinol. 159:159-170(2000).

DR	U57347; AAB39513..3;	-	DT	01-OCT-2000 (TREMBLrel. 15, Created)
DR	HSSP; P08047; 1SP2;	-	DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DR	InterPro; IPR00822; Znf-C2H2.	-	DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DR	Pfam; zf-C2H2; 3.	-	DE	PUTATIVE LIPOPROTEIN.
DR	SMART; SM00355; Znf-C2H2; 3.	-	GN	GN
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1;	3.	OS	Streptomyces coelicolor.
DR	PROSITE; PS50157; ZINC_FINGER_C2H2_2;	4.	OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
KW	DNA-binding; Metal-binding; Zinc-finger.	-	OC	Actinomycetales; Streptomycineae; Streptomyces.
FT	NON_TER 1	1	OX	NCBI_TaxID=1902;
FT	NON_TER 237	237	RN	OX
SQ	SEQUENCE 237 AA; 25421 MW; C6950DB42912DAB6 CRC64;	0;	RN	[1]
Query Match	Best Local Similarity 91.7%; Score 22; DB 6; Length 237;	0;	RP	SEQUENCE FROM N.A.
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	0;	RC	STRAINA3 (2);	
Qy 1 gtxxps 6	0;	RC	STRAINA3 (2);	
Db 204 GTATPS 209	0;	RA	Seeger K.J., Harris D.;	
Query Match	Best Local Similarity 66.7%; Score 22; DB 6; Length 237;	0;	RA	Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	0;	RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
Qy 1 gtxxps 6	0;	RN	[2]	
Db 204 GTATPS 209	0;	RP	SEQUENCE FROM N.A.	
RESULT 43	SEQUENCE FROM N.A.	0;	RC	STRAINA3 (2);
Q9L203	PRELIMINARY;	0;	RC	STRAINA3 (2);
ID Q9L203	PRT; 247 AA.	0;	RC	STRAINA3 (2);
AC Q9L203	01-OCT-2000 (TREMBLrel. 15, Created)	0;	RC	STRAINA3 (2);
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	0;	RA	Redenbach M., Kieser H.M., Denapaitre D., Eichner A., Cullum J.,	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	0;	RA	Kinashi H., Hopwood D.A.;	
DE PUTATIVE BIOTIN SYNTHASE.	0;	RT	"A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome."	
GN SCE4_05C.	0;	RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.	
OS Streptomyces coelicolor.	0;	RL	RL Mol. Microbiol. 21:77-96(1996).	
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;	0;	DR	DR EMBL; AL158061; CAB76553.1; -.	
OC Actinomycetales; Streptomycineae; Streptomyces.	0;	KW	KW Lipoprotein.	
[1]	0;	SQ	SQ SEQUENCE 253 AA; 26600 MW; FECCC3325A465D891 CRC64;	
RP SEQUENCE FROM N.A.	0;	Query Match	Best Local Similarity 91.7%; Score 22; DB 2; Length 237;	
RA Seeger K.J., Harris D.;	0;	Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	0;	
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	0;	Qy 1 gtxxps 6	0;	
RN [2]	0;	Db 122 GTASPS 127	0;	
RP SEQUENCE FROM N.A.	0;	RESULT 45	0;	
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;	0;	Q43563	0;	
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	0;	ID Q43563	PRELIMINARY; PRT; 253 AA.	
RN [3]	0;	AC Q43563;	0;	
RP SEQUENCE FROM N.A.	0;	DT 01-JUN-1998 (TREMBLrel. 06, Created)	0;	
RC STRAINA3 (2);	0;	DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)	0;	
RX MEDLINE-97000351; PubMed-8843436;	0;	DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)	0;	
RA Redenbach M., Kieser H.M., Denapaitre D., Eichner A., Cullum J.,	0;	DE ORCTL25 HYPOTHETICAL PROTEIN.	0;	
RA Kinashi H., Hopwood D.A.;	0;	GN ORCTL25_OR_BWR1B.	0;	
RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome."	0;	OS Homo sapiens (Human).	0;	
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.	0;	OC Eukaryota; Metazoa.	0;	
RL Mol. Microbiol. 21:77-96(1996).	0;	OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	0;	
DR IPR003184; CAB71805.1; -.	0;	OX NCBI_TaxID=9606;	0;	
DR InterPro; IPR003184; BioY.	0;	RN [1]	0;	
DR Pf02632; BLOY; 1.	0;	RP SEQUENCE FROM N.A.	0;	
DR SEQUENCE 247 AA; 24640 MW; 3A5EF9FDD7CF0CE CRC64;	0;	RX Cooper P.R., Smilinich N.J., Day C.D., Novak N.J., Reid L.H.,	0;	
RA Pearsall R.S., Reece M., Prawitz D., Landers J., Housman D.E., Shows T.B.,	0;	RA Schwienbacher C., Sabbioni S., Campi M., Veronese A., Bernardi G.,	0;	
RA Winterpach A., Zabel B.U., Peltier J., Weissman B.E., Shows T.B.,	0;	RA Menegatti A., Hatada T., Mukai T., Ohashi H., Barbanti-Brodano G.,	0;	
RA Higgins M.J.;	0;	RA Croce C.M., Negrini M.;	0;	
RA Genomics 0:0-0(1998).	0;	RT "Transcriptional map of 170-kb region at chromosome 11p15.5:	0;	
RN [2]	0;	RT Identification and mutational analysis of the BWR1A gene reveals the presence of mutations in tumor samples."	0;	
RP SEQUENCE FROM N.A.	0;	RT PLoS. Natl. Acad. Sci. U.S.A. 95:3873-3878(1999).	0;	
RX MEDLINE-98188297; PubMed-9520460;	0;	RL AF037066; AAC04789.1; -.	0;	
RA Pearsall R.S., Reece M., Prawitz D., Landers J., Housman D.E., Shows T.B.,	0;	DR EMBL; AF037066; AAC04789.1; -.	0;	
RA Winterpach A., Zabel B.U., Peltier J., Weissman B.E., Shows T.B.,	0;		0;	
RA Higgins M.J.;	0;		0;	
RA Genomics 0:0-0(1998).	0;		0;	
RN [2]	0;		0;	
Qy 1 gtxxps 6	0;		0;	
Db 150 GTAAAPS 155	0;		0;	
RESULT 44	SEQUENCE 247 AA; 24640 MW; 3A5EF9FDD7CF0CE CRC64;	0;		
Q9L126	PRELIMINARY;	0;		
ID Q9L126	PRT; 253 AA.	0;		
AC Q9L126;		0;		

DR	EMBL; AF035407; AAC17497.1; -	SEQUENCE	253 AA; 27218 MW; 183DE7B5C8A2255B CRC64;
Query Match	91.7%; Score 22; DB 4; Length 253;	Best Local Similarity	66.7%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Qy 1 gtxxps 6	Q91170	PRELIMINARY; PRT; 254 AA.
Db 240 GSTATPS 245	AC 091170; 01-MAR-2001 (TREMBLrel. 16, Created) DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update) PROBABLE THIOESTERASE. PA2411.	AC	091170; 01-MAR-2001 (TREMBLrel. 16, Last sequence update) PROBABLE THIOESTERASE. PA2411.
OS Pseudomonas aeruginosa.	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaeae; Pseudomonas.	OS	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaeae; Pseudomonas.
OX [1]	SEQUENCE FROM N.A.	OX	SEQUENCE FROM N.A.
RP STRAIN=ATCC 15692 / PAO1;	STRAIN=ATCC 15692 / PAO1; / Published=10984043;	RC	STRAIN=ATCC 15692 / PAO1; / Published=10984043;
RA Stover C.K.; Pham X.-Q.T.; Erwin A.L.; Mizoguchi S.D.; Warrener P.';	Stover C.K.; Pham X.-Q.T.; Erwin A.L.; Mizoguchi S.D.; Warrener P.';	RA	Stover C.K.; Pham X.-Q.T.; Erwin A.L.; Mizoguchi S.D.; Warrener P.';
RA Hickey M.J.; Brinkman F.S.L.; Hulinagle W.O.; Kowalik D.J.; LaRou M.,	Hickey M.J.; Brinkman F.S.L.; Hulinagle W.O.; Kowalik D.J.; LaRou M.,	RA	Hickey M.J.; Brinkman F.S.L.; Hulinagle W.O.; Kowalik D.J.; LaRou M.,
RA Garber R.L.; Goinkin L.; Tolentino E.; Westbrock-Wadman S.,	Garber R.L.; Goinkin L.; Tolentino E.; Westbrock-Wadman S.,	RA	Garber R.L.; Goinkin L.; Tolentino E.; Westbrock-Wadman S.,
RA Brody L.L.; Coulter S.N.; Folger K.R.; Kas A.; Larbig K.; Lim R.M.,	Brody L.L.; Coulter S.N.; Folger K.R.; Kas A.; Larbig K.; Lim R.M.,	RA	Brody L.L.; Coulter S.N.; Folger K.R.; Kas A.; Larbig K.; Lim R.M.,
RA Smith K.A.; Spencer D.H.; Wong G.K.-S.; Wu Z.; Paulsen I.T.,	Smith K.A.; Spencer D.H.; Wong G.K.-S.; Wu Z.; Paulsen I.T.,	RA	Smith K.A.; Spencer D.H.; Wong G.K.-S.; Wu Z.; Paulsen I.T.,
RA Reizer J.; Sauer M.H.; Hancock R.E.W.; Lory S.; Olson M.V.,	Reizer J.; Sauer M.H.; Hancock R.E.W.; Lory S.; Olson M.V.,	RA	Reizer J.; Sauer M.H.; Hancock R.E.W.; Lory S.; Olson M.V.,
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.";	"Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.";	RT	"Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.";
RL Nature 406:959-964 (2000).	RL Nature 406:959-964 (2000).	RL	Nature 406:959-964 (2000).
DR InterPro; IPR001031; Thioesterase.	DR InterPro; IPR001031; Thioesterase.	DR	InterPro; IPR001031; Thioesterase.
DR Pfam; PF000975; Thioesterase.	DR Pfam; PF000975; Thioesterase.	DR	Pfam; PF000975; Thioesterase.
DR Complete proteome.	DR Complete proteome.	DR	Complete proteome.
KW Sequence 254 AA;	Sequence 27878 MW; 460C521723BD7403 CRC64;	KW	Sequence 27878 MW; 460C521723BD7403 CRC64;
Query Match	91.7%; Score 22; DB 16; Length 254;	Best Local Similarity	66.7%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Qy 1 gtxxps 6	Q98TY7	PRELIMINARY; PRT; 275 AA.
Db 108 GTAAPS 113	AC 098TY7; 01-JUN-2001 (TREMBLrel. 17, Created) DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update) DE MYOD.	AC	098TY7; 01-JUN-2001 (TREMBLrel. 17, Created) DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update) DE MYOD.
OS Brachydanio rerio (zebrafish) (Zebra danio).	Brachydanio rerio (zebrafish) (Zebra danio).	OS	Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;	OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;	OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
NCBI_TaxID=7955;	NCBI_TaxID=7955;	NCBI_TaxID=7955;	NCBI_TaxID=7955;
RN [1]	SEQUENCE FROM N.A.	RN	SEQUENCE FROM N.A.
RP Gao J.; Du S.;	RP Gao J.; Du S.;	RP	Gao J.; Du S.;
RA Isolation and Characterization of Muscle-Specific Expression of	RA Isolation and Characterization of Muscle-Specific Expression of	RA	Isolation and Characterization of Muscle-Specific Expression of
RT 01-MAY-2000 (TREMBLrel. 13, Created)	RT 01-MAY-2000 (TREMBLrel. 13, Created)	RT	01-MAY-2000 (TREMBLrel. 13, Created)
Q9RN9	PRELIMINARY;	Q9RN9	PRELIMINARY;
TD 09RN9	PRT;	TD 09RN9	PRT;
AC Q9RN9;	AC Q9RN9;	AC Q9RN9;	AC Q9RN9;
DT 01-MAY-2000	DT 01-MAY-2000	DT 01-MAY-2000	DT 01-MAY-2000
RT	288 AA.	RT	288 AA.

01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE VIR99 HOMOLOG.
 GN *Bartonella henselae* (Rochalimaea *henselae*).
 OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bartonellaceae; *Bartonella*.
 OX NCBI_TaxID=38323;
 RN [1]
 RP SEQUENCE FROM N.A.
 STRAIN=HOUSTON_1;
 RX MEDLINE=20338084; PubMed=10882236;
 RA Padmala Yam L., Karem K., Baumstark B., Massung R.;
 "The gene encoding the 17-kDa antigen of *Bartonella henselae* is
 located within a cluster of genes homologous to the virB virulence
 operon.";
 RT DNA Cell Biol. 19:377-382 (2000).
 DR AF182718; AAF00947.1;
 SEQUENCE 288 AA; 32120 MW; B94F15C58F38D4B9 CRC64;

Query Match 91.7%; Score 22; DB 12; Length 288;
 Best Local Similarity 66.7%; Pred. No. 3.6e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 gtxxps 6
 DB 271 GRTSPS 276

RESULT 50
 Q9DWD6
 ID Q9DWD6 PRELIMINARY; PRT; 288 AA.
 AC Q9DWD6;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 PR58;
 DE
 GN
 OS Rat Cytomegalovirus (strain Maasricht).
 OC Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Marmegalovirus.
 OX NCBI_TaxID=79700;
 RN [1]
 RP SEQUENCE OF 1-168 FROM N.A.
 RX STRAIN=MAASRICHHT; MEDLINE=963335691; PubMed=8757999;
 RA Beuken E., Sloboe R., Bruggeman C.A., Vink C.;
 "Cloning and sequence analysis of the genes encoding DNA polymerase,
 glycoprotein B, ICP18.5 and major DNA-binding protein of rat
 cytomegalovirus.";
 RT DNA replication and functional characterisation of the origin of lytic-phase
 RT J. Gen. Virol. 77:1559-1562 (1996).
 RN [2]
 RP SEQUENCE OF 167-288 FROM N.A.
 RX STRAIN=MAASRICHHT; MEDLINE=98031207; PubMed=9367384;
 RA Vink C., Beuken E., Bruggeman C.A.;
 "Cloning and functional characterisation of the origin of lytic-phase
 DNA replication of rat cytomegalovirus.";
 RL J. Gen. Virol. 78:2963-2973 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAASRICHHT; MEDLINE=10366325; PubMed=10906222;
 RX Vink C., Beuken E., Bruggeman C.A.;
 RT "Complete DNA sequence of the rat cytomegalovirus genome.";
 RL J. Virol. 74:7656-7665 (2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAASRICHHT; MEDLINE=20473137; PubMed=11018281;
 RX Gruijthuijsen V.K., Beuken E., Bruggeman C.A., Vink C.;
 RA "Rat cytomegalovirus R89 is a highly conserved gene which expresses a

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)